

AJS

13

CUL - H05442 - B-12704

American

Journal

Sociology

Volume 103 Number 4

January 1998

EDWARD O. LAUMANN, Editor

MARY ELLEN KONIECZNY, JIM DAVIS, PAMELA POPIELARZ (University of Illinois—Chicago), MICHAEL J. ROSENFELD, and KAZUO YAMAGUCHI, Associate Editors

MARTIN RIESEBRODT, Associate Editor, Book Review

KELLY CHONG, ELIZABETH J. MCSWEENEY, DAWNE MOON, MONICA PRASAD, CHRISTOPHER B. SWANSON, and GENEVIÈVE ZUBRZYCKI, Book Review Editorial Board

SUSAN ALLAN, Managing Editor

SHARON LIU and PATRICIA L. MORSE, Editorial Assistants

Consulting Editors ROBERT J. ANTONIO • WILLIAM G. AXINN • DENISE BIELBY • LAWRENCE BOBO • KENNETH BOLLEN • PHILLIP BONACICH • BRUCE CARRUTHERS • AVSHALOM CASPI • ELISABETH CLEMENS • GERALD DAVIS • DORIS R. ENTWISLE • GARY ALAN FINE • JOHN LIE • IVAN LIGHT • BRUCE LINK • LYN H. LOFLAND • RUTH MILKMAN • MARTINA MORRIS • EUGENE ROSA • KIM LANE SCHEPPELE • JOHN SKVORETZ • AAGE SØRENSEN • ANNEMETTE SØRENSEN • ROGER WALDINGER • JONATHAN ZEITLIN • XUEGUANG ZHUO

International Consulting Editors ROSEMARY GARTNER (Canada) • JOHN HAGAN (Canada) • HIROSHI ISHIDA (Japan)

Editorial Board ANDREW D. ABBOTT • GARY S. BECKER • CHARLES E. BIDWELL • DONALD J. BOGUE • ANTHONY BRYK • MARY C. BRINTON • RONALD BURT • TERRY N. CLARK • JOHN L. COMAROFF • ROGER V. GOULD • LARRY HEDGES • EVELYN KITAGAWA • DONALD N. LEVINE • WILLIAM L. PARISH • LESLIE SALZINGER • ROBERT J. SAMPSON • ROSS M. STOLZENBERG • FRED L. STRODTBECK • GERALD D. SUTTLES • RICHARD TAUB • LINDA J. WAITE • DINGXIN ZHAO

AMERICAN JOURNAL OF SOCIOLOGY (ISSN 0002-9602) is published bimonthly (every other month) in July, September, November, January, March, and May by the University of Chicago Press, Journals Division, 5720 S. Woodlawn Ave., Chicago, Illinois 60637. Subscription rates per year: U.S. institutions, \$131.00, individuals, \$43.00, students, \$30.00 (photocopy of valid student ID must accompany subscription), ASA and BSA members, \$38.00. All other countries add \$12.00 to cover postage. Canadian subscribers add 7% GST. Subscription agent for Japan: Kinokuniya Company, Ltd. Single copy rates: institutions \$21.85, individuals \$7.15. Volumes 1-94 available from Periodical Services Co., 11 Main St., Germantown, New York 12526. All volumes are available from University Microfilms International, 300 North Zeeb Road, Ann Arbor, Michigan 48106. Subscriptions are payable in advance and will begin with the first issue published after receipt of order. Checks should be made payable to *American Journal of Sociology*. For subscription information, write University of Chicago Press, Journals Division, P.O. Box 37005, Chicago, Illinois 60637, fax, 312-753-0811; or E-mail subscriptions@journals.uchicago.edu (editorial correspondence, E-mail AJS@ccrco.uchicago.edu)

Postmaster: Send address changes to *American Journal of Sociology*, University of Chicago Press, Journals Division, P.O. Box 37005, Chicago, Illinois 60637

Copying beyond Fair Use. The code on the first page of an article in this journal indicates the copyright owner's consent that copies of the article may be made beyond those permitted by Section 107 or 108 of the U.S. Copyright Law provided that copies are made only for personal or internal use or for the personal or internal use of specific clients and provided that the copier pay the stated per-copy fee through the Copyright Clearance Center, Inc. Operations Center, 27 Congress St., Salem, Massachusetts 01970. To request permission for other kinds of copying, such as copying for general distribution, for advertising or promotional purposes, for creating new collective works, or for resale, kindly write to Permissions Department, University of Chicago Press, 5801 South Ellis Avenue, Chicago, Illinois 60637. If no code appears on the first page of an article, permission to reprint may be obtained only from the author.

Periodicals postage paid at Chicago, Illinois, and at additional mailing offices.

© 1998 by The University of Chicago. All rights reserved.

This paper used in this publication meets the minimum requirements of American National Standard for Information Sciences—Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984. (∞)

AJS

American Journal of Sociology

Volume 103 No. 4 January 1998

CONTENTS

- 837 The Determinants of Deadly Force: A Structural Analysis of Police Violence
DAVID JACOBS AND ROBERT M. O'BRIEN
- 863 Lost in Space: The Geography of Corporate Interlocking Directorates
CLIFFORD KONO, DONALD PALMER, ROGER FRIEDLAND, AND
MATTHEW ZAFONTE
- 912 Caveat Emptor: The Construction of Nonprofit Consumer Watch-
dog Organizations
HAYAGREEVA RAO
- 962 What Is Agency?
MUSTAFA EMIRBAYER AND ANN MISCHKE
- 1024 The Long-Term Effects of Foreign Investment Dependence on
Economic Growth, 1940-1990
JEFFREY KENTOR

Commentary and Debate

- 1047 Response to Hirsch's Review Essay
W. RICHARD SCOTT
- 1049 Rejoinder to Scott
PAUL HIRSCH

Review Essay

- 1052 An Intellectual's Stock in the Factory of Mexico's Ruins
CLAUDIO LOMNITZ

Book Reviews

- 1066 *The Creativity of Action* by Haris Joas
COLIN CAMPBELL

13



- 1068 *Fin de Siècle Social Theory: Relativism, Reduction, and the Problem of Reason* by Jeffrey C. Alexander
MICHÈLE LAMONT
- 1069 *Sociological Reasoning: Towards a Post-Modern Sociology* by Rob Stones
WILLIAM OUTHWAITE
- 1071 *The Social Construction of Nature: A Sociology of Ecological Enlightenment* by Klaus Eder
STEVEN NEY
- 1073 *Marginalized in the Middle* by Alan Wolfe
RICHARD MADSEN
- 1075 *Genealogies for the Present in Cultural Anthropology* by Bruce M. Knauft
MICHAEL M. J. FISCHER
- 1077 *Communicative Interaction, Power, and the State: A Method* by Frank M. Stark
JAMES FERNANDEZ
- 1078 *Rational Choice Theory and Religion: Summary and Assessment*, edited by Lawrence A. Young
JOHN R. HALL
- 1081 *Congregation and Community* by Nancy Tatom Ammerman
RHYS H. WILLIAMS
- 1083 *Religion in the Age of Decline: Organization and Experience in Industrial Yorkshire, 1870-1920* by S. J. D. Green
STEVE BRUCE
- 1084 *Glorious Battle: The Cultural Politics of Victorian Anglo-Catholicism* by John Shelton Reed
GENE BURNS
- 1086 *Radical Religion in America: Millenarian Movements from the Far Right to the Children of Noah* by Jeffrey Kaplan
FRANK LECHNER
- 1088 *Transnational Religion and Fading States*, edited by Susanne Hoerber-Rudolph and James Piscatori
DAVID A. SMILDE
- 1089 *The New Temperance: The American Obsession with Sin and Vice* by David Wagner
JOHN C. BURNHAM

- 1092 *Echoes of the Past, Epitaphs of Dissent: A South Korean Social Movement* by Nancy Abelmann
HYUN OK PARK
- 1094 *The Abolition of Feudalism: Peasants, Lords, and Legislators in the French Revolution* by John Markoff
DAVID ZARET
- 1096 *The Rhetoric of Moral Protest: Public Campaigns, Celebrity Endorsement, and Political Mobilization* by Christian Lahusen
MAYER N. ZALD
- 1098 *Crisis and Transformation: The Kibbutz at Century's End* by Eliezer Ben-Rafael
BENJAMIN D. ZABLOCKI
- 1099 *Historic Preservation: Collective Memory and Historical Identity* by Diane Barthel
PATRICK HUTTON
- 1101 *The Spectacle of History: Speech, Text, and Memory at the Iran-Contra Hearings* by Michael Lynch and David Bogen
JANE H. HILL
- 1103 *Between Politics and Reason: The Drug Legalization Debate* by Erich Goode
JAMES A. INCIARDI
- 1104 *Hybrids of Modernity: Anthropology, the Nation State and the Universal Exhibition* by Penelope Harvey
VICTORIA D. ALEXANDER
- 1106 *Interaction and Grammar*, edited by Elinor Ochs, Emanuel A. Schegloff, and Sandra A. Thompson
ALLEN D. GRIMSHAW
- 1108 *French Intellectual Nobility: Institutional and Symbolic Transformations in the Post-Sartrean Era* by Niilo Kauppi
CHARLES LEMERT
- 1110 *Fanon's Dialectic of Experience* by Ato Sekyi-Otu
DAVID SCOTT
- 1113 *Modernity and the State: East, West* by Claus Offe
FRED BLOCK
- 1115 *Realignments in the Welfare State: Health Policy in the United States, Britain, and Canada* by Mary Ruggie
TERRY BOYCHUK

- 1117 *Big Business, Strong State: Collusion and Conflict in South Korean Development, 1960-1990* by Eun Mee Kim
CHANGWON LEE
- 1119 *Hungary's Negotiated Revolution: Economic Reform, Social Change, and Political Succession, 1957-1990* by Rudolf L. Tökés
CAMILLE M. BUSETTE
- 1121 *Desert Capitalism: Maquiladoras in North America's Western Industrial Corridor* by Kathryn Kopinak
LESLIE SKLAIR
- 1122 *Ending a Career in the Auto Industry: "30 and Out"* by Melissa A. Hardy, Lawrence Hazelrigg, and Jill Quadagno
ANGELA M. O'RAND
- 1125 *Living Rooms as Factories: Class, Gender, and the Satellite Factory System in Taiwan* by Ping-Chun Hsiung
KUANG-CHI CHANG
- 1127 *Heroic Defeats: The Politics of Job Loss* by Miriam Golden
CHRIS TILLY
- 1127 *When Work Disappears: The World of the New Urban Poor* by William Julius Wilson
JOE R. FRAGIN
- 1131 *The Origins of the Urban Crisis: Race and Inequality in Postwar Detroit* by Thomas J. Sugrue
BOBBY M. WILSON
- 1132 *It Takes a Nation: A New Agenda for Fighting Poverty* by Rebecca M. Blank
NANCY A. DENTON
- 1134 *Saving our Children from Poverty: What the United States Can Learn from France* by Barbara R. Bergmann
SUZANNE M. BIANCHI
- 1136 *Development, Change, and Gender in Cairo: A View from the Household*, edited by Diane Singerman and Homa Hoodfar
VALERIE J. HOFFMAN
- 1138 *Latinos in New York: Communities in Transition*, edited by Gabriel Haslip-Viera and Sherrie L. Bayer
HÉCTOR CORDERO-GUZMÁN
- 1140 *Life Lines: Community, Family, and Assimilation among Asian Indian Immigrants* by Jean Bacon
NAZLI KIBRIA

- 1141 *Metropolitics: A Regional Agenda for Community and Stability* by Myron Orfield
ALLEN J. SCOTT
- 1144 *Gangsters: Fifty Years of Madness, Drugs, and Death on the Streets of America* by Lewis Yablonsky
DANIEL MONTI
- 1145 *Continuity and Discontinuity in Criminal Careers* by Paul E. Tracy and Kimberly Kempf-Leonard
ELMAR G. M. WEITEKAMP
- 1147 *The Self in the Family: A Classification of Personality, Criminality, and Psychopathology* by Luciano L'Abate, with the collaboration of Margaret S. Baggett
KENNETH C. LAND
- 1149 *Homicide, Race, and Justice in the American West, 1880-1920* by Clare V. McKanna, Jr.
RICHARD BLOCK
- 1151 *Screening the Los Angeles "Riots": Race, Seeing, and Resistance* by Darnell M. Hunt
RICHARD M. MERELMAN
- 1153 *White Lies: Race, Class, Gender, and Sexuality in White Supremacist Discourse* by Jessie Daniels
NOEL A. CAZENAVE
- 1154 *Teen Mothers and the Revolving Welfare Door* by Kathleen Mullan Harris
SANDRA K. DANZIGER
- 1156 *Immigration and the Family: Research and Policy in U.S. Immigrants*, edited by Alan Booth, Ann C. Crouter, and Nancy Landale
NANCY FONER
- 1158 *The Modernization of Fatherhood: A Social and Political History* by Ralph LaRossa
E. ANTHONY ROTUNDO
- 1160 *A History of the Family*: vol. 1, *Distant Words, Ancient Worlds*; vol. 2, *The Impact of Modernity*, edited by André Burguière, Christiane Klapisch-Zuber, Martine Segalen, and Françoise Zonabend
ROSANNA HERTZ

IN THIS ISSUE

DAVID JACOBS, professor of sociology at Ohio State University, does research in the politics of social control, issues in political economy, and social organization. Current projects include additional studies of the politics of outcomes in the criminal justice system and longitudinal studies of political and legal change.

ROBERT O'BRIEN is professor of sociology and department head at the University of Oregon. His research interests focus on criminology, measurement, and statistics. This article grows out of his and David Jacobs's shared research interests in macrolevel approaches to studying crime and social control.

CLIFFORD KONO is living in Tokyo, where he is teaching English and learning Japanese.

DONALD PALMER is professor of management and sociology at the University of California, Davis. He is an associate editor of *Administrative Science Quarterly*. His current research focuses on corporate acquisitions, corporate crime, executive succession, and the transformation of the American health care industry.

ROGER FRIEDLAND is professor of sociology and religious studies at the University of California, Santa Barbara. He is studying the relation between sacrality and sovereignty in Jerusalem with Richard Hecht, the politics of aesthetic transformation in the case of residential modernism in the United States with Harold Zellman, and the determinants of headquarters colocation in the United States and Italy with Paolo Gardinali.

MATTHEW ZAFONTE is a Ph.D. candidate in the Graduate Group in Ecology at the University of California, Davis. His current research focuses on the application of models from mathematical ecology to policy change and learning, spatial analysis of diffusion of innovation, and political coordination.

HAYAGREEVA RAO is associate professor of organization and management and adjunct associate professor of sociology at Emory University. His research analyzes the institutional and ecological antecedents of organizational change.

MUSTAFA EMIRBAYER is assistant professor of sociology and historical studies at the New School for Social Research. He is currently completing two book manuscripts: "Relational Pragmatics," a theoretical study, and "Democratic Education in America, 1830-1990," a historical study of the transformation of moral and civic education in American public schools.

ANN MISCHE is a graduate student in sociology at the New School for Social Research and a visiting scholar at the Lazarsfeld Center for the Social Sciences at Columbia University. She is currently completing a dissertation on the formation of civic culture across Brazilian youth activist networks.

JEFFREY KENTOR is a Ph.D. candidate in sociology at Johns Hopkins University. He is currently writing his dissertation, which examines the relationship between economic and military power and their relative effects of national mobility, 1870–1990. Other current research interests include examining the impact of foreign investment on the environment, and the relationship between inequality and urbanization.

CLAUDIO LOMNITZ is professor of history at the University of Chicago. His works include *Exits from the Labyrinth: Culture and Ideology in Mexican National Space* (University of California Press, 1992) and *Deep Mexico, Silent Mexico: The Cultural Geography of Modernity* (University of Minnesota Press, in press).

Information for Contributors

Editorial Procedures: All papers deemed appropriate for the *AJS* are sent out anonymously to readers. The editors rely heavily on the judgments of those readers, although they are not bound by them. To protect anonymity, only the title should appear on the manuscript. Attach a cover page with the title of the manuscript, your name, and your affiliation. All identifying references and footnotes should appear on a separate page. Manuscripts are accepted for publication subject to the understanding that the *AJS* has the right of first publication. Submission of a manuscript to the *AJS* is taken to indicate the author's commitment to publish in this *Journal*. No paper known to be under consideration by any other journal will be reviewed by the *AJS*. In addition, the *AJS* is not a suitable vehicle for materials published elsewhere, in whole or in part. Upon publication, all rights, including subsidiary rights, are owned by the University of Chicago Press. The author retains the right to use his or her article without charge in any book of which he or she is the author or editor after it has appeared in the *Journal*.

Manuscript Acceptance Policy: While it is our policy to require the assignment of copyright on most journal articles and review essays, we do not usually request assignment of copyright for other contributions. Although the copyright to such a contribution may remain with the author, it is understood that, in return for publication, the *Journal* has the nonexclusive right to publish the contribution and the continuing right, without limit, to include the contribution as part of any reprinting of the issue and/or volume of the *Journal* in which the contribution first appeared by any means and in any format, including computer-assisted storage and readout, in which the issue and/or volume may be reproduced by the publisher or by its licensed agencies.

Preparation of Copy

1. Type *all* copy—including indented matter, footnotes, tables, and references—*double spaced*, allowing generous margins at the top, bottom, and sides of page. Please do not break words at ends of lines or justify right-hand margins. Do not use proportional spacing fonts. Indicate italics by underlining only. Number all footnotes, beginning with the acknowledgment note as number 1. *Provide a word count on the cover page of the manuscript.*
2. Begin each table on a separate page. Refer to each table in numerical order in the text. Tables should not contain vertical lines or more than 20 two-digit columns or the equivalent.
3. Clarify all mathematical symbols (e.g., Greek letters) with words in the margins of the manuscript.
4. Original or glossy prints of professionally drawn figures will be required if manuscript is accepted.
5. Include a brief abstract (not more than 100 words) summarizing the findings.
6. Submit *five copies* of the manuscript and abstract, on white paper. Manuscripts should be addressed to the Editor of the AMERICAN JOURNAL OF SOCIOLOGY, 5835 South Kimbark, Chicago, Illinois 60637. *Manuscripts are not returned to authors.* Enclose a self-addressed stamped postcard to obtain acknowledgment of manuscript receipt. No other acknowledgment will be sent.
7. Each article submitted to the *AJS* must be accompanied by a check or money order for \$15.00, payable to the University of Chicago Press in U.S. currency or

its equivalent. Papers will not be reviewed until this fee has been paid. The submission fee will be waived for papers solely by student authors. Submissions from students must be accompanied by a faculty member's certification of their student status. Citizens of countries with restrictions on the export of U.S. dollars may request waiver of the submission fee.

Format of References in Text: All references to monographs, articles, and statistical sources are to be identified at the appropriate point in the text by last name of author, year of publication, and pagination where appropriate, all within parentheses. Footnotes are to be used only for substantive observations. Specify subsequent citations of the same source in the same way as the first one; do not use "ibid.," "op. cit.," or "loc. cit."

1. When author's name is in the text: Duncan (1959). When author's name is not in text: (Gouldner 1963).

2. Pagination follows year of publication: (Lipset 1964, pp. 61-65)

3. For more than three authors, use "et al." For institutional authorship, supply minimum identification from the beginning of the complete citation: (U.S. Bureau of the Census 1963, p. 117).

4. With more than one reference to an author in the same year, distinguish them by use of letters (a, b) attached to the year of publication: (1965a).

5. Enclose a series of references with a single pair of parentheses, separated by semicolons: (Duncan 1959; Gouldner 1963; Lipset 1964, pp. 61-65).

6. Format for references to machine-readable data files is:

Davis, James Allan. 1978. GSS (General Social Survey) NORC ed. Chicago: NORC. Distributed by Roper Public Opinion Research Center, New Haven, Conn

Format of References: List all items, including machine-readable data files (MRDF), alphabetically by author (providing the full list of multiple authors) and, within author(s), by year of publication, in an appendix titled "REFERENCES." For examples, see text of articles in this issue.

(Rev. 1/93)

The Determinants of Deadly Force: A Structural Analysis of Police Violence¹

David Jacobs
Ohio State University

Robert M. O'Brien
University of Oregon

Political or threat explanations for the state's use of internal violence suggest that killings committed by the police should be greatest in stratified jurisdictions with more minorities. Additional political effects such as race of the city's mayor or reform political arrangements are examined. The level of interpersonal violence the police encounter and other problems in departmental environments should account for these killing rates as well. Tobit analyses of 170 cities show that racial inequality explains police killings. Interpersonal violence measured by the murder rate also accounts for this use of lethal force. Separate analyses of police killings of blacks show that cities with more blacks and a recent growth in the black population have higher police killing rates of blacks, but the presence of a black mayor reduces these killings. Such findings support latent and direct political explanations for the internal use of lethal force to preserve order.

What factors determine the amount of violence used by democratic states when they try to control their citizens? State coercion, particularly in the most developed societies, is rarely studied by sociologists, yet it may be a powerful explanation for social order. As Weber pointed out long ago, the

¹ We gratefully acknowledge financial support from the Harry Frank Guggenheim Foundation. We thank Pam Jackson and Gary LaFree for their comments on prior drafts and Stephen E. Haynes, Joe Stone, and James Ziliak for their advice. David Schlosberg and Scott Harris deserve our thanks for helping us collect the data. The data utilized in this research were made available in part by the Inter-University Consortium for Political and Social Research. The data from the Uniform Crime Reports, 1976–1986: Supplementary Homicide Reports were originally collected by the Federal Bureau of Investigation. Neither the consortium nor the collector bear any responsibility for the results or the interpretations presented here. Direct comments to David Jacobs, Department of Sociology, 300 Bricker Hall, 190 North Oval Mall, Ohio State University, Columbus, Ohio 43210-1358

crucial defining element of the state is its ability to call upon superior force in any dispute with internal or external rivals. Without this supremacy governments cease to exist. Even contemporary democratic states frequently command scarce resources from unwilling subjects with at least the threat of violence. Although modern democratic states avoid the military and employ domestic agencies like the police to control internal miscreants and dissidents, the primary activity that makes police organizations distinct is their use of force (Bittner 1990).

Perhaps what appears as a consensus within modern societies is deceiving because much of the domestic tranquillity we so frequently observe is based on accommodation rather than consent. Sociologists often respond by asserting that a society based on pure coercion would be inefficient. Goode (1972, p. 510) agrees, but answers that "sociologists have felt that force was a weak reed to lean on and no regime or society can rest on force alone. Indeed it cannot, but neither must it do so. None ever need try; rulers and conquerors . . . always command other resources." According to Goode, scholars who use the pure coercion argument to dispute the importance of force as a prerequisite for order have rejected an empty claim.

Although violence by social control agencies in advanced states is comparatively unusual, that does not make it ineffective. We may have been misled by the relative absence of conspicuous force in these societies. In any case, the discipline has neglected coercion and its chilling effects on nonconforming behavior, particularly in the most advanced industrial democracies. While we do not believe that coercion or its threat is the only important explanation for order, we suspect that further study of the conditions that lead to state violence will increase our understanding of how order is maintained in advanced societies.

Developments in historical political sociology have established the importance of political violence, but theoretical advances in the sociology of law and social control based on the work of Vold (1958), Dahrendorf (1959), Turk (1969), and Black (1976) furnish some of the more useful explanations for this outcome. Sociologists who stress coercive explanations for order often see the control agencies of the state as principally serving the interests of the privileged. According to this political threat view, a primary (but not the only) use of criminal law and law enforcement agencies is to maintain control over the "dangerous classes" who threaten public order.

Many Weberian and most neo-Marxist theorists who see coercion as a major way to preserve order think that differences in rewards are based on power (Lenski 1966; Collins 1975). According to this view, the privileged benefit greatly from existing arrangements while many citizens receive much less; inequality is an unstable condition that must be sustained by

sanctions or their threat. If racial or economic disparities are at least partly sustained by coercion, enhanced state violence can be expected in areas where economic disparity and the menace of an underclass are greatest. We study such political explanations (and many others) for variation in the amount of internal state violence by looking at rates of lethal force within political subunits. This study assesses the determinants of the rate of killings committed by the police in 170 U.S. cities.

The conventional assumption is that police violence is a reaction to the violence the police encounter, so killings by police officers are a necessary response to the brutality they must control. A political approach suggests instead that police violence should be especially likely in areas where racial or economic divisions with political consequences are severe. If the privileged are threatened by the disorderly potential of a racial or economic underclass with little to lose and much to gain from redistributive violence, diminished efforts to control the police or demands that they be unfettered should be likely in racially or economically divided cities. Political explanations suggest that police killings will be most common in economically stratified cities with larger percentages of minorities because dominant groups have much to lose from threats to public order by a racial or economic underclass.

In this article we see if the police use of lethal force can be explained by political arrangements or by social divisions that are likely to have political consequences, but we also assess reactive accounts that stress interpersonal violence and other problems officers encounter in the urban environments they must police. Because race is such an important consideration in a study of police violence and because our initial results suggest that the determinants of police killings of blacks deserve separate study, we analyze the factors that explain these race-specific lethal force rates as well. Results from studies that assess many competing hypotheses are more accurate (Johnston 1984), so we use an inclusive strategy. This means the theoretical section that follows cannot focus on just a few explanations.

SOME TESTABLE HYPOTHESES

The Empirical Literature

Studies of the effects of racial and economic stratification on social control have examined the size of police departments. Some find minority threat to be the strongest explanation. Research by Liska, Lawrence, and Benson (1981) and Huff and Stahure (1980) on the per capita number of police, or Jackson's (1989) research on expenditures, shows that the percentage of blacks in a city is related to police strength. An analysis of metropolitan statistical areas (MSAs) found that, while percentage black mattered, eco-

nomic inequality was a stronger explanation for department size (Jacobs 1979). Also, inequality explains shifts in police strength over time (see Jacobs and Helms 1997). Disputes in this literature about the relative importance of inequality or minority presence have not been resolved.

The statistical studies of police killings are problematic. Kania and Mackey (1977) computed killings by the police in U.S. states. They ignore minorities and use welfare recipients to measure inequality, although the states with the greatest economic inequality had the smallest welfare roles (Dye 1969). Such problems make Kania and Mackey's rank order correlations too misleading to report. A reanalysis of the Kania and Mackey data (Jacobs and Britt 1979) using multiple regression finds that income inequality is the strongest explanation for police killings, but state-level data are not the ideal for this work, as they inappropriately combine many kinds of police departments that operate in disparate environments.

We use disaggregated city data (for a discussion of bias caused by aggregation, see Thiel [1971]), an imperfect unit of analysis that has advantages. This unit lets us treat police killings as an organizational outcome and analyze the environmental determinants of the lethal behavior of many individual departments.

Sherman and Langworthy (1979) and Liska and Yu (1992) appear to be the only multivariate city-level investigations. Both studies measured police killings by combining data from *Vital Statistics* with surveys of police departments, but the *Vital Statistics* estimates are about half as large as the estimates from either survey. The shared variance in the merged data that formed the dependent variables is extremely modest in both studies. In Sherman and Langworthy's study of 32 cities, the correlation between the estimates from their survey and the *Vital Statistics* estimates of police killings is .56; in Liska and Yu this correlation is .45. The problematic construction of both dependent variables plus the small number of cases and other difficulties make the inconsistent findings of these two investigations too questionable to report.

The problems identified in other studies suggest that an analysis of police killing rates in 170 cities based on a single comprehensive data source should yield more accurate results for the following reasons:

1. Information from disparate sources need not be combined to calculate the number of these lethal events.
2. Killings by the police need not be averaged across disparate departments that operate in extremely dissimilar environments because we analyze city rather than state data.
3. The large number of cases should produce superior statistical inferences and allow for the control of more explanations than those in the earlier research.

4. Local conditions, made invisible by state-level data, can now be examined.

Finally, in contrast to past studies, we use an estimator designed to handle the censored dependent variables produced by using rates computed on these infrequent events.

Alternative Ways to Account for Police Killings

The theories we assess in this article fall into two broad categories: political explanations and reactive explanations. We begin by discussing *political threat explanations*. Killings by police officers may be especially likely where divisions based on racial or economic differences are greater because such divisions threaten dominant groups and undercut effective community controls over the police.

Second, reactive hypotheses that have nothing to do with social or political divisions may explain police violence. Reactive accounts suggest that the use of lethal force by the police should be most likely in areas where the police must control a violent population or where they must react to urban conditions that make their work difficult. We begin our theoretical discussion by talking about hypotheses that focus on social divisions that have political implications and hypotheses that are explicitly political. We finish by discussing some plausible reactive explanations for the use of lethal force by the police.

THREAT THEORIES

Race and Police Violence

The menace of a racial or economic underclass with an interest in redistributive violence may lead to violent law enforcement. Blalock (1967) argues that dominant populations will be threatened in areas with a larger racial underclass. Liska, Lawrence, and Sanchirico (1982) find that, with the crime rate held constant, fear of crime is associated with the percentage of African-Americans in cities. Departments should be more likely to use lethal force in cities with greater percentages of minorities because the threat posed by a large racial underclass may lead to harsh law enforcement measures. Threat theories also suggest that law enforcement should be more coercive in cities that have recently experienced growth in the percentage of black residents, because as the population of African-Americans increases, concerns about the prevalence of crime increase as well.

This and other versions of the threat hypothesis are plausible ways to

explain at least some killings committed by the police as long as a disavowal is kept firmly in mind. It is *not* necessary to claim that privileged groups make direct demands that the police use greater amounts of deadly force. Without restraints, police violence is probable (Chevigny 1995). Officers are expected to preserve order rather than enforce the law (Bittner 1990), and violence is an efficient (but unfortunate) way to achieve this end. Police behavior is not readily observed, yet extreme personal risks are an inseparable part of this work. These personal risks give officers reasons to use violence when it may not be warranted. Where the perceived threat of underclass violence is greater, all that is required for political explanations for police killings to hold is for the powerful to be less willing to interfere with police methods. Hughes (1963) argues that a willingness to remain ignorant about the activities of the "dirty workers" is a compelling explanation for much official brutality. Dominant groups in cities with greater racial or economic disparities who are threatened by these conditions *may* demand that the police be unfettered, but the political threat explanations tested in this article do not require such active participation. Political divisions would still explain police killings if privileged groups only are less likely to interfere with police violence in cities where they feel threatened by an underclass.²

The use of lethal force may be associated with political arrangements as well. Wilson (1971) claims that cities with progressive "good government" reforms have police departments that are comparatively insulated from political restraints because reform administrative arrangements are less responsive to majority pressure than more conventional political arrangements. This insulation may reduce community control over police violence, so cities with city manager or commission governments may have higher police killing rates.³

We test a second explicitly political hypothesis by looking at the race of the mayor. Police chiefs, who often originate and always enforce departmental policies that regulate officer proclivities to use force, serve at the

² Hence, external institutionalized controls on the police may be less effective in jurisdictions where privileged groups who could restrain the police use of deadly force choose to remain ignorant about the methods these "dirty workers" use to maintain order. Unfortunately, we cannot find systematic information about the presence of external review boards.

³ Chevigny (1995) and others claim that officers in Los Angeles were likely to use violence because civil service policies made the removal of the chief extremely difficult. The public and their elected representatives in this exceptional city could not control police violence because they could not remove the primary police administrator. Sampson and Cohen (1988) find that active police departments that make greater numbers of arrests are likely to be found in cities with reform governments. It is reasonable to believe that the more aggressive officers in these cities also may be more likely to use lethal force.

option of the mayor in almost all jurisdictions. If the mayor is black and probably dependent on minority votes, police administrators should have greater incentives to curb police violence. This condition should make them particularly likely to revise internal policies to achieve this goal. In our data, blacks are the victims of 53% of all killings by police, so the race of the mayor should matter.

Ethnographic and other work on the police shows that departmental regulations exercise strong controls on officer behavior (Rubenstein 1973; Fyfe 1982; Sherman 1983). It follows that police killings should be reduced in cities with a black mayor because the officer on the street should realize that violence against African-Americans will be more closely investigated and perhaps more severely sanctioned than it will be in cities with white mayors.

Inequality and Police Violence

Following Jacobs and Britt (1979), economic stratification should explain the amount of deadly force used by the police. The economic version of the political perspective on state violence holds that disparities in economic rewards produce a potentially unstable social order that must be maintained by force. Chambliss and Seidman (1980, p. 33) write that "the more economically stratified a society becomes, the more it becomes necessary for dominant groups to enforce through coercion the norms of conduct that guarantee their supremacy." Many neo-Marxists believe that, without the threat of coercion, unequal arrangements cannot be sustained.

Neo-Marxists do not have a monopoly on this view. Turk (1966, p. 349) observes that when there are greater differences in resources "normal legal procedures are likely to be unofficially abrogated in favor of summary and less costly procedures" because procedural law is expensive. For all of these reasons we expect that cities with greater economic differences between the affluent and the poor should have higher rates of killing by the police.

Theorists who stress coercion often see inequality as a potentially unstable condition that must be sustained by force or by its threat. Many of these theorists also hold that economic resources are a source of power. Because power is a relational asset that rests on differences in resources, inequality is unlikely to give the lower class or minorities greater control over the police for several reasons. First, because they have less power, people with the least resources should have a reduced ability to protect themselves from police violence in unequal areas. Second, if racial or economic inequality must be sustained by force, privileged groups in unequal cities will have less reason to interfere with police violence, and third, in unequal areas the privileged may indirectly encourage police killings

by demanding that restraints be removed so the police can maintain order.

Yet hypotheses that focus only on economic inequality ignore the relative economic position of minorities. Such explanations assume that economic divisions between either poor whites or poor blacks and the affluent have identical effects on social control agencies. We overcome this problem by looking at the effects of economic disparities between blacks and whites. In cities where the ratio of black to white mean incomes most favors whites, the relative number of police killings should be greater for several reasons.

First, the greater the differences between black and white economic resources, the more a relatively poor black underclass with less to lose from redistributive violence should threaten privileged whites who, in this situation, may be less willing to curb police violence. Second, both neo-Marxists and sociologists who are not Marxists (e.g., Blalock 1967) view power as at least partly based on differences in economic resources. Substantial differences in the economic resources of blacks and whites should reduce the black population's political influence and thus their control over police violence. Finally, racial inequality may lead threatened whites to make direct demands that the police be unfettered to deal with the threat posed by a potentially violent racial underclass.

Hence theorists who stress coercion often see inequality as a potentially unstable condition that must be sustained by force or by its threat. Many of these theorists also hold that economic resources are a source of power. Because power is a relational asset that rests on differences in resources, inequality is unlikely to give the lower class or minorities greater control over the police for several reasons. First, because they have less power, people with the least resources should have a reduced ability to protect themselves from police violence in unequal areas. Second, if racial or economic inequality must be sustained by force, privileged groups in unequal cities will have less reason to interfere with police violence, and third, in unequal areas the privileged may indirectly encourage police killings by demanding that restraints be removed so the police can maintain order.

In his review of field studies that examined the use of lethal force, Sherman (1980b) concludes that about half of the people fired on by the police did not have guns. The percentage shot while fleeing also was substantial. He claims that these "executions without trial occur in response to crimes against property without any defense justification" (p. 89). Sherman's conclusion that such deaths often result from overreaction leaves much variance for political (and other) hypotheses to explain after violence in departmental environments is held constant.

REACTIVE HYPOTHESES

Civilian Violence and Violent Police Responses

Some police killings undoubtedly are a reaction to difficult conditions in departmental environments. If findings about social divisions or direct political hypotheses are to be credible, it is crucial to include extensive controls for these problematic urban conditions. For example, the police should be especially likely to use violent methods where they must deal with a violent population. This most fundamental reactive explanation suggests that police departments in cities with higher civilian murder rates should be more likely to use deadly force.

Family Strife and Police Violence

Sampson (1987) and others find that broken families lead to violent crime. We assess this effect in two ways. First, police killing rates may be greater in cities with larger percentages of female-headed families. Second, because much police work involves coping with violent domestic disputes and because these interventions often require force (Rubenstein 1973), police departments in cities with substantial divorce rates should be more likely to use lethal violence. For these reasons, we use both divorce rates and broken families to measure the intrafamily strife that may lead to more killings by police officers.

Population Density and Police Violence

Poverty or crowding, measured by dwelling units with more than a threshold number of people per room, may produce violent behavior that leads to violent responses by the police. Deadly force also may be likely in jurisdictions where departments have fewer officers relative to the population they must control. Finally, the police encounter heightened difficulties in larger cities in part because anonymity is enhanced. This relationship and the effectiveness of inexpensive informal controls in smaller cities suggests that greater amounts of deadly force can be expected in the most populous jurisdictions.

A SUMMARY OF THE HYPOTHESES

Political threat theories.—Police may use deadly force because they protect the interests of the privileged by keeping the redistributive violence of subordinate racial or economic groups in check. If the *economic* version is correct, the most economically unequal cities will

have the most police killings. If the *racial* version of this political threat approach is correct, the following four relationships should hold.

- (1) Police killing rates will be higher in cities with more blacks or
- (2) in cities where growth in the percentage of black residents has been pronounced.
- (3) We can expect a greater use of lethal force where there are larger differences in the economic resources of blacks and whites.
- (4) The police should not use lethal force as often in cities with a black mayor.

Reactive hypotheses.—The police may kill because they must protect all groups from violence. If reactions to civilian violence explain police violence, we can expect more police killings in the cities with the most murders or in cities that are especially difficult to police for other reasons, including family strife, poverty, and crowding. But, according to the reactive view, law-abiding citizens benefit from the public safety provided by this seemingly necessary use of lethal violence.

METHODOLOGY

Research Design, Dependent Variable, and Estimation

Because the environmental approach has been so productive, most empirical studies of organizations stress environmental explanations. Sherman (1986a) applies this logic to the police behavior when he says, "Theoretically the community level should be given the most attention. . . . Rossi [Berk, and Edison] (1974) found that in comparing city of employment and officer's personal characteristics as explanations of the use of aggressive detection tactics, 67 percent of the variance was uniquely attributable to the city" (p. 94).

Environmental factors ought to explain a police department's propensity to use lethal force as well, so we analyze the rate of police killings in 1980 in the 170 U.S. cities with a population greater than 100,000. The data come from supplementary homicide reports that police departments filed with the FBI. This information is likely to be accurate because such homicides are difficult to conceal. The alternative information on police killings relies on the erratic behavior of coroners and medical examiners (Sherman and Langworthy 1979).

Police killings are unusual events.⁴ We therefore compute the total rate

⁴ The mean number of *total* police killings across all 170 cities, 1980–86 is 9.68 compared to a mean of 1.99 for the *rates* per 100,000. The five cities with the highest rates are Fort Lauderdale, 7.41; Dallas, 7.30; St. Louis, 7.17; Newark, 7.02; and New Orleans, 6.83. It is interesting that Los Angeles ranked twentieth with a police killing rate of 4.23, placing this city in the eighty-eighth percentile.

of police killings per 100,000 residents in a city over seven years starting in 1980 and ending at 1986.⁵ This aggregation is not novel (see Kania and Mackey 1977; Jacobs and Britt 1979; Liska and Yu 1992), but after computing these averages, 31 cities (18.3%) still had no police killings. Because ordinary least squares (OLS) gives inconsistent estimates when the dependent variable has many zero values, Tobit frequently is employed to deal with such censored dependent variables.

Tobit uses two formulas to predict values of the dependent variable—one for cases at the limit value (zero in this study) and another for cases above the limit (Roncek 1992, p. 503), so this procedure jointly calculates two effects. Tobit combines probit estimates of the probability that a case will have a nonlimit value for cases at the limit on the dependent variable with estimates of the effects of the explanatory variables on values of the dependent variable for cases with a nonlimit value on the dependent variable. We capitalize on the properties of this estimator because it is the most appropriate when the dependent variables are censored rates with large numbers of zeros (Roncek 1992; Greene 1993).⁶

We provide separate analyses of the rate of blacks killed by the police in the same period, but we must deal with an added methodological difficulty. Cities with a tiny black population can have extremely high rates of police killings of blacks. If a city has a small number of blacks, enormous rates can result from just one or two incidents over the entire seven-year period. To remove this distortion, we weight the Tobit equations explaining police killings of blacks by the number of blacks in a city. This procedure substantially reduces the effects of cities with few blacks and appropriately increases the influence of cities with a large number of black

⁵ If we add three years and compute mean police killings from 1980 to 1989, the theoretical implications of the findings do not change, but we are dubious about such long lags. In any case, using these additional years to reduce the number of zeros in the dependent variable is ineffective. For example, 28 zeros remain in total police killing rates if police killings are averaged from 1980 to 1989 compared to 31 when the last year used is 1986. For these reasons, we handle zeros in the dependent variable by using Tobit. We do *not* average data for the regressors across multiple years because some theoretically important independent variables (e.g. Gini or the black/white income ratio) are available only for 1979. Averaging an independent variable over many years almost invariably increases that variable's explanatory power. Because multiple-year averages cannot be computed for every independent variable, they *must not* be used for any. We use city means for missing killings in 1982 for Houston and in 1984 and 1985 for Chicago, but these changes have infinitesimal effects on the results.

⁶ For additional discussion and references to the many studies that have used Tobit, see Greene (1993) or any other standard econometrics text. For precedent in the use of this procedure by sociologists, see Walton and Ragin (1990), Roncek and Maier (1990), and Mosher and Hagan (1994). We use the Tobit routine in STATA ver 5.0.

residents.⁷ To correct skewed distributions, "1" is added to dependent variables and they are transformed into natural log form.

Measurement of Explanatory Variables

We measure minority presence with the percentage of blacks in a city (%BLACK). Squaring this variable weights the highest %BLACK values more heavily and puts greater emphasis on the largest black populations that should be most threatening (for precedent, see Jacobs and Helms [1996]). This transformation captures a crucial part of the threat hypotheses we are testing. We assess the effects of shifts in %BLACK with percentage change in %BLACK (indicated as %CHANGE) as measured from 1970 to 1980. To correct for a modestly skewed distribution, this variable is in square root form.⁸

In supplemental (but unreported) analyses, we use the percentage of nonwhite population instead, but this measure gives almost identical results. If the separate combined percentage of white Hispanics and other racial minorities who are not black is added to equations that already contain the %BLACK variable, the nonwhite variable is insignificant. For these reasons, we measure minority effects with data on blacks.

We operationalize racial inequality with the ratio of black-to-white mean family incomes (ratios of medians give similar results) while aggregate inequality is measured with a Gini index calculated on household incomes as reported by the 1980 census. We include a dummy scored "1" for the presence of black mayors and a dummy for location in the South.⁹ We measure poverty with the percentage of all families below the poverty line or with the percentage of impoverished black families. Following Sampson (1987), breakdowns in the family are assessed with the percentage of the population that is divorced or by the percentage of female-

⁷ When we analyze police killings of blacks, we therefore follow econometric precedent and weight by the number of blacks to remove the effects of heteroscedasticity. Such weights appropriately reduce the effects of cities with few blacks and with extremely high police killing rates. Seventy cases, or 41.2% of the 170 cities, had no police killings of blacks, but weighting by the number of blacks substantially reduces the effects of these cases with zero values.

⁸ Some variables are in square root or natural log form to correct for skewed distributions. These transformations reduce the effects of outliers and increase the likelihood of bivariate normality (Fox 1991).

⁹ Dayton, Ohio, shifted from a black to a white mayor in the early 1980s during the years when yearly police killings were combined. If we alter the value of this case from "1" to "0" to reflect this shift, the results remain virtually identical. Unfortunately, we could not find data on the race of police chiefs or the racial makeup of departments in 1979 or 1980.

headed black families. Crowding is measured with the natural log of the percentage of housing units with more than 1.01 residents per room.

We assess police strength with the number of police employees per 100,000 residents. We use total number of employees because departments substitute unsworn personnel for uniformed officers, but this choice does not matter. We measure violence in departmental environments with the murders per 100,000 residents.¹⁰ Attempts to use violent crime rates or the robbery rates indicate that the murder rates are the best control. Murder rates (and the other explanatory variables in these 170 cities) are highly correlated from one year to the next, so we find that using either 1979 or 1980 rates furnishes almost identical results.

In the analyses of the police killings of blacks, we use the percentage of black families with a female head (because this variable has stronger effects than the black divorce rate), the black murder rate, which is computed by dividing the number of black offenders by the black population, and the percentage of black families who have incomes below the poverty line. Explanatory variables are measured in 1980.

Model Specification

One general specification of the Tobit model that predicts total police killings is

$$\begin{aligned} \text{POLKILL}_i = & b_0 + b_1\% \text{BLACK}_i^2 + b_2\% \text{CHANGEB}_i \\ & + b_3 \text{BLK/WHITEINC}_i + b_4 \text{MURDRT}_i \\ & + b_5 \text{POP}_i + b_6 \text{DIVRC}_i \\ & + b_7 \text{BLKMAYOR}_i + b_8 \text{SOUTH}_i \\ & + b_9 \text{CITYGOV}_i + b_{10} \text{CROWDING}_i \\ & + b_{11} \text{GINI}_i \text{ (or } b_{11} \text{POV}_i + \text{RESIDUAL}_i \end{aligned} \quad (1)$$

where POLKILL is the natural log of the rate of police killings per 100,000 citizens plus one, %BLACK is the square of the percentage of blacks in the population, %CHANGEB is the square root of the percentage change in the percentage of blacks in the population, BLK/WHITEINC is black mean family income divided by white mean family income, MURDRT is the number of murders per 100,000 population, CITYGOV is a dummy scored "1" if a city has a city manager or a commission form of govern-

¹⁰ Deaths caused by the police are removed from all independent variables that measure killings. Killings by police officers are eliminated from both the total murder rate and the black murder rate.

ment, POP is the natural log of population, DIVRC is the percentage of divorced persons in the population, BLKMAYOR is a dummy variable for a black mayor, SOUTH is a dummy for the South, CROWDING is the natural log of the percentage of housing units with more than 1.01 persons per room, GINI is the index of income inequality, and POV is the percentage of families with incomes below the poverty line.

Because they are too collinear to be used in the same equation, GINI and POV are given the same subscript and separated by *or*. In additional equations we enter the number of police employees per 100,000 population. Only the coefficients on the black/white income ratio, black mayors, and police employees per capita should be negative.

A general equation for the weighted (by the number of blacks) Tobit model predicting police killings of blacks is

$$\begin{aligned} \text{POLKILLB}_i = & b_0 + b_1\% \text{BLACK}_i + b_2\% \text{CHANGEB}_i \\ & + b_3 \text{BLK/WHITEINC}_i + b_4 \text{BLKMURDRT}_i \\ & + b_5 \text{POP}_i + b_6 \text{BLKFMLHD}_i \\ & + b_7 \text{BLKMAYOR}_i + b_8 \text{SOUTH}_i \\ & + b_9 \text{CITYGOV}_i + b_{10} \text{CROWDING}_i \\ & + b_{11} \text{GINI}_i \text{ (or) } b_{11} \text{POV}_i + \text{RESIDUAL}_{i1} \end{aligned} \quad (2)$$

where POLKILLB is the natural log of the police killing rate of blacks per 100,000 blacks plus one, BLKMURDRT is the black murder rate, BLKFMLHD is the percentage of black families headed by a female, and the remaining variables are defined as above. Again we expect that only the black/white income ratio, black mayors, and police employees per capita should be inversely related to police killing rates of blacks.

ANALYSIS

Table 1 gives means and standard deviations for the variables. There is substantial variation in the dependent variable. Seven-year mean police killing rates across these 170 cities range between zero and 7.41 police killings per 100,000 residents.

Tobit Results

Total killings.—Table 2 presents the Tobit estimates of total police killing rates. Equation (1) in table 2 shows a restricted six-variable model. In equation (2) we add the black mayor measure and the dummy for location in the South. In equation (3) we add police employees per 100,000, reform

TABLE 1
MEANS AND SDs OF VARIABLES

	Mean	SD
POLKILL,	921	.602
BLK/WHITEINC,676	.134
%BLACK ¹	631 460	940 786
%CHANGEB	10.422	4 456
DIVRC	10 035	2 457
MURDRT	14.524	11 029
Police employees per 100,000 population	259 347	96.746
POP	12.291	752
GINI	368	.036
BLKMAYOR065	247
CITYGOV	568	497
CROWDING	1.431	.538
SOUTH	361	.462
POV	11.066	4.788

NOTE.—Computed on 170 US cities with population over 100,000. Except for the income data used to compute inequality scores, explanatory variables are based on 1960 data.

political arrangements (CITYGOV), and the measure of income inequality (Gini). In equation (4) we drop police strength and add crowding. In equation (5) we substitute percent below the poverty line for Gini in an equation otherwise identical to equation (4).

These results show that murder rates, population, and divorce rates are consistent predictors of the use of lethal force by the police, but other environmental conditions that make police work more difficult do not matter.¹¹ Neither crowding, poverty, nor the relative size of departments explain this outcome. Some indicators of political divisions also do not predict these deaths. In contrast to Jacobs and Britt's (1979) results using state data, economic inequality is unrelated to police killings. Other racial and political variables that do not appear to influence the total rate of police killings include %BLACK, %CHANGEB, CITYGOV, and the presence of a black mayor.

Regardless of how the specifications are altered, the same four explana-

¹¹ Substituting the robbery rate or the violent crime rate in a specification otherwise the same as that used in eq. (4) of table 2 leads to identical theoretical implications, but the explanatory power is reduced. The robbery rates and the murder rates are too collinear to be used together, but forcing them into the same analysis does not alter the theoretical conclusions. The evidence strongly suggests that murder rates are the best indicator of the primary reactive explanation for police killing rates.

TABLE 2

TOBIT ANALYSES OF ALL KILLINGS BY POLICE OFFICERS

	(1)	(2)	(3)	(4)	(5)
Intercept ..	-1.9636* (-2.31)	-1.9897* (-2.27)	-2.2758 (-1.69)	-2.5965 (-1.76)	-2.5481** (-2.76)
%BLACK ¹ ..	-0.0568 (-.09)	-0.0110 (-.16)	-0.0187 (-.27)	.0021 (.03)	-.0257 (-.35)
%CHANGE ..	.0076 (.74)	.0076 (.74)	.0087 (.82)	.0071 (.65)	.0108 (1.00)
BLK/WHITEINC ..	-1.4921*** (-3.72)	-1.4692*** (-3.33)	-1.4181** (-3.02)	-1.4001** (-2.77)	-1.5341*** (-3.39)
MURDRT ..	.0262*** (4.58)	.0261*** (4.50)	.0252*** (4.15)	.0241*** (3.72)	.0210*** (3.13)
POP ..	.2278*** (3.65)	.2291*** (3.63)	.2309*** (3.56)	.2411*** (3.72)	.2459*** (3.81)
DIVRC ..	.0580*** (3.16)	.0573*** (3.01)	.0534** (2.63)	.0529** (2.52)	.0661** (2.89)
BLKMAVOR ..		.0292 (.14)	.0232 (.11)	.0147 (.70)	.0221 (.11)
SOUTH ..		.0148 (.14)	...	-.0060 (-.06)	-.0280 (-.27)
CITYGOV ..			.1373 (1.46)	.1141 (1.22)	.1329 (1.40)
Police per 100,000 population ..			.0006 (1.16)
GINI ..			.1057 (.96)	.3240 (.16)	..
CROWDING0671 (.74)	.0326 (.35)
POV0226 (1.31)
Log likelihood ..	-138.55	-138.53	-137.05	-137.44	-136.59
χ^2 ..	95.89***	95.93***	98.89***	98.11***	99.81***

NOTE.—%CHANGE is in square root form while %BLACK¹ has been divided by 10,000. Nos. in parentheses are *t*-values.* $P \leq .05$; ** $P \leq .01$; *** $P \leq .001$.

tory variables account for variation in police killings. The results suggest that the police use of deadly force is greatest in the most populated cities and in cities with the highest murder rates. Officers in communities with larger divorce rates also are more likely to use lethal violence, but the findings indicate that an important social division that is likely to have political consequences explains this outcome as well. Police killings are especially likely in cities where economic differences between blacks and whites are most pronounced.

Because the racial inequality measure is theoretically important and such a consistent predictor of the total rate of police killings and because blacks are the victims of 53% of all of the police killings in our sample, we next present separate analyses of the police use of lethal force against African-Americans.

Police killings of African-Americans.—Table 3 shows the results of the Tobit analyses of the use of lethal force against blacks. The explanatory variables are similar to those in table 2, but we substitute the black murder rate (BLKMURDRT) for the total murder rate and the black poverty rate (BLKPOV) for the total poverty rate. Because the black divorce rate does not explain police killings of blacks, we use the percentage of black families headed by a female instead. All models again are estimated with Tobit. In these analyses, however, the equations are weighted by the number of blacks in a city to correct the heteroscedasticity that results from a few cases with extremely high black victimization rates that is largely an effect of extremely small black populations in some cities.¹²

The specifications follow the pattern employed in table 2. Equation (1) in table 3 presents another restricted model limited to six explanatory variables. In equation (2) we again add the presence of a black mayor and the dummy for location in the South. In equation (3) we add police employees per 100,000, reform political arrangements, and the aggregate measure of income inequality (Gini). In equation (4) we drop police strength, and add crowding. In equation (5) we substitute the percentage of black families below the poverty line for Gini in an equation otherwise identical to equation (4).¹³

¹² Cities with the highest scores on police killings of blacks per 100,000 blacks are Torrance, Calif., 228.65; Sunnyvale, Calif., 39.22; Amarillo, Texas, 22.48; Long Beach, Calif., 21.18; Oklahoma City, 19.22; Arlington, Texas, 18.87, and Jacksonville, Fla., 17.45. Both Torrance and Sunnyvale have fewer than 1,000 black residents and just one (Sunnyvale) or two (Torrance) police killings between 1980 and 1986. This means that a weighting scheme like the one we use that reduces the influence of these two cases is appropriate. The mean *number* of police killings of blacks across these seven years is 5.13. The mean *rate* of police killings of blacks is 5.94 per 100,000 blacks.

¹³ The unweighted means and SDs for variables introduced in the analysis of police killings of blacks are as follows: the mean of the natural log of police killing rates of

The explanatory variables that mattered when all police killings are analyzed continue to explain police killings of blacks despite the weighting and other changes. Blacks are more likely to be killed by the police in larger cities with higher black murder rates and more broken families, although the measure of broken families is different in these equations. Unreported findings show that the percentage of all female-headed families does not explain the total rate of police killings, but the results in table 3 show that the percentage of black families headed by a female is positively associated with police killings of blacks. The coefficients on these race-specific variables provide additional support for reactive explanations.

Some theoretically interesting contrasts are revealed by comparing the coefficients on the measures of political divisions and the coefficients on the direct measures of political effects in tables 2 and 3. The racial inequality variable is significant in both analyses, but the presence or absence of a black mayor does not influence total police killings. When the analysis is confined to the use of lethal force against blacks, however, we find that police killings are likely to be reduced when the most powerful political official in a city is an African-American. Finally, when we analyze all police killings, neither the percentage of blacks nor the percentage change in the percentage of the population that is black is significant. Yet when the police killing rates of blacks are at issue, both of these racial threat measures predict these deaths.

ALTERNATIVE TESTS

Because these findings provide insight about such an important outcome, we conducted sensitivity tests. The effects of additional explanatory variables were assessed in unreported equations. Residential segregation is not associated with either dependent variable. Unemployment rates, unemployment disaggregated by race, change in unemployment rates, either the rate or the number of police officers who were killed, and the presence of young males disaggregated by race have the same negligible effects on the total rates of police killings and police killings of blacks.¹⁴ Using different combinations of the explanatory variables also does not alter the results. Finally, we could find no evidence for departures from linearity.

blacks plus one is 1.22 (SD 1.13); the mean of the black murder rate is 38.11 (SD 30.14); the mean of the percentage of female-headed black families is 27.49 (SD 7.05).

¹⁴ Collinearity is not problematic in any of these analyses. All VIF scores are well below the threshold value of 10 that conservative statisticians advocate as an indicator of collinearity. This outcome and the stability of the results despite the many diverse specifications support a conclusion that this difficulty is not distorting our findings.

TABLE 3

WEIGHTED TOBIT ANALYSES OF POLICE KILLINGS OF BLACKS

	(1)	(2)	(3)	(4)	(5)
Intercept	-1.0300 (-1.50)	-2.0636* (-2.22)	-8489 (-.60)	-1.4832 (-.98)	-1.9560 (-1.88)
%BLACK'	.1015** (2.93)	.1798*** (3.51)	.1752** (2.50)	.1855*** (3.27)	.1802*** (3.27)
%CHANGE	.0680* (1.71)	.0707* (1.79)	.0791* (1.94)	.0680* (1.69)	.0695* (1.73)
BLK/WHITEINC	-1.6968*** (-3.54)	-1.5364** (-2.78)	-1.9731*** (-3.25)	-1.6803** (-2.62)	-1.5559** (-2.71)
BLKMURDR	.0160*** (5.79)	.0192*** (5.95)	.0202*** (6.23)	.0193*** (5.82)	.0194*** (5.75)
POP	.1219*** (3.20)	.1456*** (3.29)	.0988* (1.79)	.1495** (2.55)	.1416* (2.26)
BLKFMLHD	.0266* (2.20)	.0379** (2.90)	.0288* (2.00)	.0402** (2.71)	.0401** (2.51)
BLKMAYOR	..	-3517* (-1.95)	-4036* (-2.26)	-3378* (-1.84)	-3627* (-1.90)
SOUTH	..	.2003 (1.39)	..	.2121 (1.44)	.2167 (1.36)
CITYGOV	-.0193 (-.14)	-.0257 (-.19)	-.0322 (-.24)
Police per 100,000 population0003 (.45)
GINI	-.3579 (-12)	-1.4271 (-45)	..
CROWDING0078 (.05)	-.0029 (-.02)
BLKPOV	-.0036 (-.24)
Log likelihood	-158.93	-155.05	-155.88	-154.93	-155.00
χ^2	60.63***	68.39***	66.74***	68.64***	68.49***

NOTE: -%CHANGE is in square root form while %BLACK' has been divided by 10,000. Nos. in parentheses are *t*-values.

* $P \leq .05$; ** $P \leq .01$; *** $P \leq .001$.

To see if weighting leads to the contrasts between the unweighted findings reported in table 2 and the weighted findings reported in table 3, we weighted the Tobit equations that predict total police killings by city population. But when these weighted results are compared to the equivalent unweighted results in table 2, the same explanatory variables are significant.

We use principal components to combine Gini, the total (and, where appropriate, the race-specific) poverty rates, and crowding to create a comprehensive measure of the presence of an underclass, but these scales have insignificant effects when their component variables are dropped and the scales are added to the equations. When we reestimate the models using White's (1980) correction for heteroscedasticity, the corrected *t*-values do not contradict the reported results.¹⁵ The Tobit equations in the tables also pass the link test (Pregibon 1980) for specification error.

In accord with the conventional wisdom, we find that measures of interpersonal violence in departmental environments and other problematic urban conditions lead to more police killings, but racial threat effects and one explicitly political indicator matter as well.

DISCUSSION

A Review of the Results

These findings are robust. When we substitute explanatory variables, the results persist. The findings also hold when various sensitivity tests are conducted. Plausible contrasts appear when different police killing rates are analyzed. One explanation for these stable findings is the comparative accuracy of supplemental homicide data on police killings. The number of cities employed in this study probably contributes to the stability of the findings as well and allows us to assess many competing explanations. According to Johnston (1984), estimates based on comprehensive models are less likely to be biased.¹⁶

¹⁵ We analyze police killing rates rather than simple counts because the likelihood of being killed by the police is theoretically more interesting. Analyses of the *number* of police killings using negative binomial regression (an estimator similar to Poisson regression but without such restrictive assumptions about the distribution of the dependent variable) nevertheless give results with the same theoretical implications as those found using Tobit. Another check for distributional problems like heteroscedasticity involves transforming the dependent variables to ranks by rounding to the nearest integer and then reestimating with ordinal probit. When this is done, the Tobit findings reported in the tables persist.

¹⁶ Johnston says, "It is more serious to omit relevant variables than to include irrelevant variables since in the former case the coefficients will be biased, the disturbance variance overestimated, and conventional inference procedures rendered invalid, while in the latter case the coefficients will be unbiased, the disturbance variance properly estimated, and the inference procedures properly estimated. This constitutes

The findings suggest that many of these lethal events stem from problematic urban conditions. The appropriate murder rates explain both total police killings and police killings of blacks. Population matters in all equations as well, so a conclusion that the police are more likely to use lethal force in the most populous cities seems warranted. This finding is plausible because anonymity and other conditions that increase the difficulties faced by the police are enhanced in larger cities.

Some of the remaining measures of problematic conditions in departmental environments do not influence the use of lethal force. Poverty rates, the number of police employees per 100,000, and crowding are unrelated to police killings. Higher divorce rates are associated with larger total police killing rates, however, and the presence of female-headed black families also predicts the police use of deadly force against blacks.

In contrast to Jacobs and Britt's (1979) state-level analysis, we find no evidence that aggregate inequality measured by the Gini index explains police killings. But studies that use cities often do not find that economic inequality predicts criminal justice outcomes. When the affluent suburbs and the central cities are combined by using data from MSAs, coefficients on economic inequality typically become stronger (Jacobs 1979). Limiting an analysis to cities excludes many prosperous residents who live in the suburbs, so community inequality probably is underestimated when researchers use city data. Yet it must be remembered that this is a post hoc explanation for a hypothesis that was ineffective.

The findings instead suggest that, at least during the 1980s, racial inequality (BLK/WHITEINC) is the kind of economic stratification that best explains the use of deadly force by the police. The results are consistent with a view that greater differences in the economic resources of blacks and whites reduce the black population's political influence and their ability to curb police violence. These findings also support a parallel hypothesis derived from threat theory that the willingness of dominant groups to interfere with harsh law enforcement methods is diminished in cities where differences in the economic resources of blacks and whites are most pronounced.

We also find that minority presence matters. While the percentage of blacks in the population is unrelated to total police killings, it has positive relationships with the use of deadly force against blacks. The latter result supports prior findings about the determinants of formal social control.

a fairly strong case for including rather than excluding relevant variables in equations. There is, however, a qualification. Adding extra variables, be they relevant or irrelevant, will lower the precision of estimation of the relevant coefficients" (1984, p. 262), so inclusive specifications typically provide more conservative significance tests

According to these findings, cities with greater numbers of blacks tend to have stronger law enforcement agencies that operate in a more punitive fashion. The standard interpretation for such results hinges on threat effects. Following Blalock (1967), threat theorists argue that, where blacks are relatively numerous, white dominance is threatened, so whites use social control agencies to maintain their ascendant position. Our finding that change in the percentage of blacks in a city's population is positively associated with police killings of blacks after other factors are controlled provides additional support for this racial threat account.

Since they are often victims of crime, many blacks may support a stronger police force or other policies that offer the hope of reducing these risks. Studies of social control that find a positive relationship between the presence of blacks and police strength or arrest rates do not offer unequivocal support for the racial threat hypothesis because many African-Americans may see such measures as desirable. In contrast to the studies that analyzed these control outcomes, this study focuses on behavior that is far more likely to hurt blacks. A comparison of racial victimization rates shows that the rate of police killings of blacks per 100,000 blacks is much greater than the the same rate for non-Hispanic whites (5.13 vs. .977). Because African-Americans have far less reason to condone the absence of restraints on police violence, an analysis of police killings offers a better test of racial threat theory than analyses of police strength or arrests.

Perhaps this disproportionate probability that blacks will be killed by the police is the main reason why the race of the mayor matters. While the relationships between the presence of reform political arrangements and police killings are negligible, our findings show that cities with a black mayor have fewer police killings of blacks. This finding is plausible because black mayors, who hire and fire police administrators, are likely to be dependent on black votes. Hence, black mayors have both the motive and the political resources to reduce police killings of blacks. These results suggest that social scientists should not overlook political factors when they analyze the determinants of police violence.

IMPLICATIONS AND CONCLUSIONS

The results reported above challenge views that focus only on economic inequality and ignore racial minorities. Arguments that economic gaps between either poor whites or poor blacks and the affluent have identical effects on repressive efforts to maintain order are not corroborated by findings that racial hypotheses explain police killings or by the weak associations between economic inequality and these lethal events, but they do show that racial inequality accounts for both the total rate of police kill-

ings and the rate of blacks killed by the police. The persistence of all of these racial results after murder rates and other difficult urban conditions have been held constant, moreover, does not fit with the conventional wisdom that the use of deadly force by the police is only a reaction to the interpersonal violence they must control.

Macro studies such as this one can isolate general relationships, but aggregate data are unlikely to furnish information about the intervening links between structural explanations and political outcomes. The links between structural arrangements that give dominant races greater influence and outcomes that are consistent with their interests often must be inferred. In this case, we have suggested plausible mechanisms that explain how racial inequality and the other structural effects isolated in this study may lead to a greater use of deadly force.

We acknowledge, however, that ideological factors enumerated by Garland (1990) or by Savelsberg (1994) may complement our structural accounts. In racially stratified cities social distance and fear of an underclass may reduce an ability to see minorities and the poor as equally deserving of protection from executions without trial. Liska et al.'s (1982) finding that there is a positive association between minority presence and fear of crime that persists after the crime rates have been held constant fits with this view, but we have no evidence about beliefs.

This study nevertheless has important implications. Sociologists have not stressed coercion as an explanation for order in advanced societies. Perhaps the ubiquity of market exchanges has directed our attention away from force or its threat (Collins 1975). Because the violence used by contemporary control agencies is modest compared to times when rulers constructed states by forcefully subduing domestic rivals, students of advanced societies neglect these critical events (Tilly 1978).

Yet those who assume that state violence is unimportant forget that coercion, if used to regulate market exchanges, is the ultimate source of control in market economies. If most people realize that visible defiance of the state's coercive supremacy will be unprofitable, that does not mean that the threat of state violence no longer is an important explanation for social order. It suggests instead that we look closely at the political processes that determine how these critical resources are employed.

According to Weber, politics is the struggle for control over the coercive resources of the state. Since the police are to the state as the edge is to a knife (Bayley 1985), political explanations for police violence should be fruitful. Yet the political determinants of the behavior of social control agencies rarely are studied because social scientists often assume that state agencies charged with using domestic coercion serve universal interests. The results of this study and those found in other studies motivated by

a political approach to law enforcement challenge this view. By uncovering the racial and economic divisions that lead to harsh law enforcement methods, a relatively new research area that might be called the *political sociology of social control* has increased our knowledge about the coercive determinants of social order in modern societies.

This study extends this tradition by assessing the environmental determinants of police killings. A finding that the police use of lethal force varies with the degree of inequality between the races, the presence of blacks, and local political arrangements that increase black control over the behavior of law enforcement personnel supports political explanations for these violent events. Such results are consistent with claims that state violence is used in racially unequal jurisdictions to preserve the existing order. They also suggest that sociologists should not neglect the coercive foundations of domestic order particularly in advanced but racially divided societies like the United States.

REFERENCES

- Bayley, David. 1985. *Patterns of Policing*. New Brunswick, N.J.: Rutgers University Press.
- Bittner, Eon. 1990. "The Police on Skid Row." Pp. 30–62 in *Aspects Police Work*, edited by Eon Bittner. Boston: Northeastern University Press.
- Black, Donald. 1976. *The Behavior of Law*. New York: Academic Press.
- Blalock, Hubert. 1967. *Towards a Theory of Minority Group Relations*. New York: Capricorn Books.
- Chambliss, William J., and R. Seidman. 1980. *Law, Order, and Power*. Reading, Mass.: Addison-Wesley.
- Chevigny, Paul. 1995. *The Edge of the Knife*. New York: New Press.
- Collins, Randall. 1975. *Conflict Sociology*. New York: Academic Press.
- Dahrendorf, Ralf. 1959. *Class and Class Conflict in Industrial Societies*. Stanford, Calif.: Stanford University Press.
- Dye, Thomas A. 1969. "Income Inequality and American State Politics." *American Political Science Review* 63:157–62.
- Fox, John. 1991. *Regression Diagnostics*. Newbury Park, Calif.: Sage Publications.
- Fyfe, James J. 1982. "Administrative Interventions on Police Shooting Discretion: An Empirical Examination." Pp. 258–81 in *Readings on Police Use of Deadly Force*, edited by James J. Fyfe. Washington, D.C.: Police Foundation.
- Garland, David. 1990. *Punishment and Modern Society*. Chicago: University of Chicago Press.
- Goode, William J. 1972. "The Place of Force in Modern Society." *American Sociological Review* 37:507–19.
- Greene, William H. 1993. *Econometric Analysis*. New York: Macmillan.
- Huff, C. Ronald, and John M. Stahure. 1980. "Police Employment and Suburban Crime." *Criminology* 17:461–70.
- Hughes, Everett C. 1963. "Good People and Dirty Work." In *The Other Side*, edited by Howard Becker. New York: Free Press.
- Jackson, Pamela L. 1989. *Minority Group Threat, Crime, and Policing*. New York: Praeger.

- Jacobs, David. 1979. "Inequality and Police Strength. Conflict Theory and Social Control in Metropolitan Areas." *American Sociological Review* 44:913-25.
- Jacobs, David, and David W. Britt. 1979. "Inequality and the Police Use of Deadly Force." *Social Problems* 26:403-12.
- Jacobs, David, and Ronald E. Helms. 1996. "Toward a Political Model of Incarceration: A Time-Series Examination of Multiple Explanations for Prison Admission Rates." *American Journal of Sociology* 102:323-57.
- . 1997. "Testing Coercive Explanations for Order: The Determinants of Law Enforcement Strength over Time." *Social Forces* 75:1361-92.
- Johnston, J. 1984. *Econometric Methods*. New York: McGraw-Hill.
- Kania, Richard, and Wade Mackey. 1977. "Police Violence as a Function of Community Characteristics." *Criminology* 15:27-48.
- Lenski, Gerhard. 1966. *Power and Privilege*. New York: McGraw-Hill.
- Liska, Allen E., J. J. Lawrence, and M. Benson. 1981. "Perspectives on the Legal Order." *American Journal of Sociology* 87:412-26.
- Liska, Allen E., J. J. Lawrence, and A. Sanchirico. 1982. "Fear of Crime as a Social Fact." *Social Forces* 60:760-71.
- Liska, Allen E., and Jiang Yu. 1992. "Specifying and Testing the Threat Hypothesis." Pp. 53-68 in *Social Threat and Social Control*, edited by Allen Liska. Albany: State University of New York Press.
- Mosher, Clayton, and John Hagan. 1994. "Constituting Crime in Upper Canada: The Sentencing of Narcotics Offenders, circa 1908-1953." *Social Forces* 72:613-41.
- Pregibon, D. 1980. "Goodness of Link Tests for Generalized Linear Models." *Applied Statistics* 20:15-24.
- Roncek, Dennis W. 1992. "Learning More from Tobit Coefficients." *American Sociological Review* 57:503-7.
- Roncek, Dennis W., and Pamela A. Maier. 1990. "Bars, Blocks, and Crimes Revisited: Linking the Theory of Routine Activities to the Empiricism of 'Hot Spots.'" *Criminology* 29:725-53.
- Rossi, Peter H., Richard Berk, and B. K. Edison. 1974. *The Roots of Urban Discontent: Public Policy, Municipal Institutions, and the Ghetto*. New York: Wiley.
- Rubinstein, Jonathan. 1973. *City Police*. New York: Farrar, Straus & Giroux.
- Sampson, Robert J. 1987. "Urban Black Violence." *American Journal of Sociology* 93:348-82.
- Sampson Robert J., and Jacqueline Cohen. 1988. "Deterrent Effects of the Police on Crime." *Law and Society Review* 22:163-89.
- Savelsberg, Joachim J. 1994. "Knowledge, Domination, and Criminal Punishment." *American Journal of Sociology* 99:911-43.
- Sherman, Lawrence W. 1980a. "Causes of Police Behavior." *Journal of Research in Crime and Delinquency* 17:69-100.
- . 1980b. "Execution without Trial: Police Homicide and the Constitution." *Vanderbilt Law Review* 33:71-100.
- . 1983. "Reducing Police Gun Use: Critical Events, Administrative Policy, and Organizational Change." Pp. 98-125 in *Control in the Police Organization*, edited by Maurice Punch. Cambridge, Mass.: MIT Press.
- Sherman Lawrence W., and Robert H. Langworthy. 1979. "Measuring Homicide by Police Officers." *Journal of Criminal Law and Criminology* 70:546-60.
- Theil, Henri. 1971. *Principles of Econometrics*. New York: Wiley.
- Tilly, Charles. 1978. *From Mobilization to Revolution*. Reading, Mass.: Addison-Wesley.
- Turk, Austen T. 1966. "Conflict and Criminality." *American Sociological Review* 31:338-52.
- . 1969. *Criminality and the Legal Order*. Chicago: Rand McNally.

- Vold, George. 1958. *Theoretical Criminology*. New York: Oxford.
- Walton, John, and Charles Ragin. 1990. "Global and National Sources of Political Protest: Third World Responses to Debt Crises." *American Sociological Review* 55: 876-90.
- White, Halbert. 1980. "A Heteroscedasticity-Consistent Covariance Matrix and a Direct Test for Heteroscedasticity." *Econometrica* 48:817-38.
- Wilson, James Q. 1971 *Varieties of Police Behavior*. New York: Atheneum

Lost in Space: The Geography of Corporate Interlocking Directorates¹

Clifford Kono

University of California, Santa Barbara

Donald Palmer

University of California, Davis

Roger Friedland

University of California, Santa Barbara

Matthew Zafonte

University of California, Davis

The article studies the causes of local and nonlocal interlocking directorates among the largest U.S. industrial corporations in 1964. The authors hypothesize that interlocks are spatial phenomena—with spatial attributes and spatial determinants. Consistent with this hypothesis, they find that local and nonlocal interlocks have different correlates. Further, three spatial structures influence interlocking: the location of a corporation's headquarters vis-à-vis other corporate headquarters and upper-class clubs, the territorial distribution of a firm's production facilities, and the spatial configuration of a corporation's ownership relations. This suggests that previous interlock research, which ignores spatial considerations, has been seriously misspecified.

INTRODUCTION

Interlocking directorates, which arise when two corporations share one or more directors, have long been suspected of eroding market competition and political democracy in the United States. Many contend interlocks

¹ We thank P. Devereaux Jennings for help in computing the spatial interlock variables used in this study and Kathy Thoren and Christina Lozano for preparation of the text, tables, and figures. We are indebted to William Bielby, Fred Block, James Lincoln, Beth Mintz, Harvey Molotch, Peter Phillips, Michael Schwartz, Barry Staw, Robert Sutton, and John Walton for comments on earlier versions of this article. We also appreciate the criticisms and suggestions of the *AJS* reviewers, who helped us refine our arguments and analyses. Special thanks to our children—Hannah, Sarah, Erica, and Tony—without whom we would be truly lost in space. Direct correspon-

facilitate the flow of information between firms (Mills 1956; Stanworth and Giddens 1975) and the transmission and enforcement of norms among business elites (Koenig and Gogel 1981; Useem 1983). Some argue that they allow corporations to co-opt or infiltrate one another and thereby coordinate their behavior (Pfeffer and Salancik 1978; Mintz and Schwartz 1985; Glasberg 1989). A few assert that interlocks, when associated with interfirm stockholding or lending relationships, even serve as vehicles of interfirm control (Fitch and Oppenheimer 1970). Recent research demonstrates that interlocking directorates influence a wide range of corporate behaviors—political action committee contributions, public policy group memberships, organizational form changes, and involvement in mergers and acquisitions (see Mizruchi [1996] for a review).

A host of theories offer competing explanations of the causes of interlocking. Most believe that interlocks are created to serve the organizational interests of corporations or the class interests of the elites who command them. On the one hand, traditional resource dependence theorists believe that interlocks are the result of corporate attempts to reduce uncertainty and constraint inherent in the market structure of the economy (Pfeffer 1972; Allen 1974; Pfeffer and Salancik 1978; Burt 1980, 1983; Burt, Christman, and Kilburn 1980). Extending this view, bank control and bank hegemony theorists believe that interlocks reflect the paramount importance of capital allocation within the economy (Fitch and Oppenheimer 1970; Kotz 1978; Mintz and Schwartz 1981a, 1981b, 1985). On the other hand, social class theorists believe that interlocks reflect the internal structure of the capitalist class, differentiated by kinship ties among major stockholders (Sweezy 1953; Zeitlin 1974; Zeitlin, Ratcliff, and Ewen 1974) or social ties among business elites (Bonacich and Domhoff 1981; Useem 1979, 1984; Soref 1976; Soref and Zeitlin 1987). However, a few still maintain that interlocks are largely unintended consequences of managerial attempts to acquire the advice of experienced business friends (Mace 1986).

Empirical support for these theories is mixed at best. While studies at the industry level of analysis indicate that nonfinancial resource dependence relations are associated with interlocking among manufacturing firms (Pfeffer and Salancik 1978; Burt 1980, 1983), firm-level analyses do not (Galaskiewicz et al. 1985; Palmer, Friedland, and Singh 1986; Mizruchi 1989, 1992). Firm-level analyses of the relationship between financial dependence and financial interlocking have produced conflicting results (Dooley 1969; Pfeffer 1972; Allen 1974; Pfeffer and Salancik 1978; Burt 1980; Pennings 1980; Mintz and Schwartz 1985; Stearns and Miz-

dence to Donald Palmer, Graduate School of Management, University of California, Davis, California 95616.

ruchi 1987; Mizruchi and Stearns 1988). Tests of the social class perspective have only been conducted at the person level of analysis (Soref 1976, 1980; Bearden and Mintz 1987; Johnson and Mintz 1989). Mace's (1986) expertise and friendship argument has never been tested at the firm level for U.S. firms. Finally, to the best of our knowledge, no one has evaluated each of these theories of interlocking in a single, multitheoretic study.

Our goal in this article is to assess competing theories of corporate interlocking by analyzing interlocking as a spatial phenomenon. Social theory tends to ignore the spatial character of social action and structure, implicitly viewing space as a stage upon which they unfold (Friedland and Boden 1994; Soja 1989). Contemporary organization and class theory are written as if corporations, their administrative and productive activities, and their leaders are not situated in a physical world.¹ The failure to theorize spatial structure and process in interlock research is exemplified by the relative silence surrounding its most established finding, that the network of interlocking directorates is localized in space—consisting of “clusters” of corporations interlocked with other firms headquartered in their headquarters locale, which are in turn tied together by “bridges” formed by corporations interlocked nationally (Dooley 1969; Levine 1972; Allen 1974, 1978; Koenig and Sonquist 1977; Mizruchi 1982; Mintz and Schwartz 1985; Bearden and Mintz 1985, 1987; Johnson and Mintz 1989). It is important that this spatial structure has been observed in studies that span the period from 1935 through 1976—a time well after the rise of jet travel and instantaneous long-distance voice and image transmission, conventionally assumed to have progressively reduced the importance of space.

Researchers have debated the extent to which region rather than other factors is the *principal* dimension along which the interlock network is segregated and have speculated about the factors—common resource exchange or financial control relations—binding firms together in regional cliques (Allen 1978; Mizruchi 1982; Mintz and Schwartz 1981a, 1981b, 1985). But, with one partial exception (Allen 1978), no one has analyzed the spatial structure of corporate interlocking at the firm level. Specifically, no one has explored the possibility that local and nonlocal interlocking have different determinants. If local and nonlocal interlocking have different determinants, past research may have underestimated, if not misconstrued, the impact of some hypothesized determinants. Nor has anyone explored the possibility that the geography of organizational and class

¹ There are exceptions to this tendency. A growing number of contemporary theorists view space as a constitutive dimension of social phenomena (Foucault 1977; Harvey 1982, 1985, 1989; Bourdieu 1990; Giddens 1990). Within organizational studies, population ecologists have begun to explore the significance of space (Barnett and Carroll 1987; Haveman and Nonnemaker 1997; Haveman and Romanelli 1997).

relationships in which firms are embedded shapes interlocking. If it does, past theory about the causes of interlocking is incomplete.

In this article, we analyze the spatial structure of interlocks maintained by large U.S. industrial corporations in 1964. We determine the extent to which corporations interlocked with other industrial firms and financial institutions and then decompose these totals into their local and nonlocal components. We then model total, local, and nonlocal industrial and financial interlocking as a function of determinants suggested by previous theories, as well as an urban growth machine theory partly of our own design.

THEORY AND HYPOTHESES

Expertise and Friendship

The simplest explanation of interlocking holds that persons are elected to a corporation's board on the basis of their business expertise and friendship with the firm's inside directors (Mace 1986). Corporate executives try to place knowledgeable outsiders on their boards because boards advise managers on long-term strategy. Because the directors of large corporations tend to possess the greatest business expertise, this creates a network of interlocking directorates among the largest corporations. Managers also try to place friends on their boards because boards can support them against threats to their tenure. Because spatial proximity is an important precondition of friendship between people (Festinger, Schacter, and Back 1950; Festinger 1953; Feld 1981), this probably leads to interlocks with locally headquartered firms. Thus, the number of other firms headquartered in a corporation's headquarters location should be directly related to the likelihood that it interlocks with other firms, especially other local firms.¹

To the best of our knowledge, no one has examined the relationship between the spatial propinquity of corporate headquarters and interlocking among U.S. firms at the firm level. Green (1983) found that levels of interlocking between firms in different U.S. and Canadian cities were inversely proportional to the distance between cities. Lincoln, Gerlach, and Takahashi (1992) report that Japanese corporations are more likely to share directors when they were headquartered in the same prefecture (a Japanese provincial unit of governance similar to a U.S. county). Ka-

¹ While professional search firms today are responsible for a growing percentage of director appointments, the majority are still arranged without the aid of formal intermediaries. In 1983, about 9% of all outside corporate directors were identified by executive search firms (Miruk and Giardina 1983). By 1989, this figure had risen—but only to 16% (Korn/Ferry International 1989).

TABLE 1

EXCLUSIVE METROPOLITAN CLUBS IDENTIFIED BY DOMHOFF (1970)
AS UPPER-CLASS INSTITUTIONS

Club	City	Club	City
Arlington	Portland	Maryland	Baltimore*
Boston	New Orleans*	Millwaukee	Millwaukee
Brook	New York	Minneapolis	Minneapolis
Burlingame Country	San Francisco	New Haven	
California	Los Angeles	Lawn	New Haven*
Casino	Chicago	Pacific Union	San Francisco
Century	New York	Philadelphia	Philadelphia
Chargin Valley Hunt	Cleveland	Piedmont Driving	Atlanta*
Charleston	Charleston*	Piping Rock	New York
Chicago	Chicago	Racquet	St. Louis
Cuyamaca	San Diego	Rainer	Seattle
Denver	Denver	Richmond	
Detroit	Detroit	German	Richmond
Eagle Lake	Houston	Rittenhouse	Philadelphia
Everglades	Palm Beach*	River	New York
Hartford	Hartford	Rolling Rock	Pittsburgh
Hope	Providence	Saturn	Buffalo
Idlewild	Dallas	St. Cecelia	Charleston*
Knickerbocker	New York	St. Louis Country	St. Louis
Links	New York	Somerset	Boston
		Union	Cleveland
		Woodhill Country	Minneapolis

* Clubs situated in cities in which none of the 1964 top 500 industrial corporations were headquartered.

dushin (1995) found that French financial elites were no more likely to sit on the same boards of directors if they resided in the same exclusive Paris neighborhood—the 16th arrondissement. However, elites who lived in the 16th arrondissement were more likely to be friends and elites who were friends were more likely to serve on the same corporate boards.

Upper-Class Social Interaction

Friendships that give rise to director appointments may be predicated on more than spatial propinquity. Inside directors prefer board candidates whom they trust will remain loyal to them in times of crisis. The 40 exclusive metropolitan upper-class social clubs that Domhoff (1970) maintains are sites of capitalist class formation (listed in table 1) may provide vehicles for the development of such trust. Domhoff (1970, 1974) and others (Useem 1987; Phillips 1994) believe that these clubs are foci for elite inter-

action that provide institutionalized informal settings in which elites are socialized and socially controlled to adhere to normative business attitudes and behaviors on which trust is predicated. Hence the local presence of upper-class clubs may facilitate friendships that lead corporate directors to select one another to serve on each others' boards. While local upper-class clubs provide a firm's directors with the opportunity to develop relationships of trust with both local and nonlocal club members, contact with locally employed directors should be most intimate and most likely to lead to board nominations. Thus, being headquartered in a city with an upper-class club should be directly related to the likelihood that corporations interlock, especially with other local firms.

No one has explored the possibility of an ecological impact of exclusive upper-class clubs on the interlocking behavior of corporations headquartered in the cities where they are situated. However, several studies analyze the connection between directors' upper-class club memberships and corporate board affiliations. Most show that directors who belong to upper-class clubs hold more corporate and financial board seats and thus create more industrial and financial interlocks than directors who do not belong to these clubs (Bonacich and Domhoff 1981; Soreff 1976, 1980; Soreff and Zeitlin 1987). Further, Johnson and Mintz (1989) demonstrate that the probability that two directors serve on two or more of the same boards is higher when both belong to an upper-class club. They also found that membership in the same upper-class club was associated with the probability that two directors served on two or more of the same corporate boards—but only when directors were principally affiliated with firms headquartered in different regions. They speculate that social ties less formal than common club memberships give rise to common corporate board memberships when directors reside in the same region—implying, in contradiction to our hypothesis above, that the effect of local upper-class clubs on interlocking will be most evident in connection with nonlocal interlocking.

Resource Dependence Relations

The traditional resource dependence perspective.—The dominant organizational explanation of interlocking assumes that corporations create interlocks in response to uncertainty and constraint stemming from their reliance on other firms for resources. Corporations experience uncertainty vis-à-vis competitors, depending on the concentration levels in the industries in which they produce (Pfeffer and Salancik 1978). Firms are constrained by buyers and suppliers, depending on the extent to which they transact with only a few other industries, as well as the extent to which their industries and those with which they transact are highly concen-

trated (Burt 1983). Corporations are constrained by financial institutions to the degree that they borrow from them (Dooley 1969; Allen 1974; Burt 1980; Pfeffer and Salancik 1978). Corporations can reduce resource uncertainty and constraint by establishing linkages with firms producing in industries that generate uncertainty and constraint for them (Pfeffer and Salancik 1978; Burt 1980). Interlocking directorates are one such linkage, potentially providing corporations with information about and influence over problematic firms in their environment. Thus, the number of interlocks corporations maintain should be associated with the level of resource uncertainty and constraint to which they are exposed.

Traditional resource dependence theorists also assume a corporation's size influences its propensity to interlock. Large corporations are thought to pose more uncertainty for other firms and thus are assumed to more often be the targets of other firms' interlock strategies (Pfeffer and Salancik 1978).⁴ A corporation's industrial diversity may also influence its propensity to interlock. On the one hand, diversification is a strategy to reduce uncertainty and constraint, reducing the amount of information and influence firms need to manage their environments, thereby possibly decreasing the need to interlock (Pfeffer and Salancik 1978). On the other hand, industrial diversity increases the complexity of corporate environments and thus increases the amount of information and influence firms need to manage them, thereby likely increasing the need to interlock. Hence, we remain equivocal about the direction of industrial diversity's effect on interlocking.

The traditional resource dependence perspective has largely ignored spatial considerations. However, in an early programmatic statement of this perspective, Dill (1958) noted that organizations operating in many spatially dispersed locales require more information to manage their environments. This implies that geographic complexity should be directly related to the number of interlocks corporations maintain. Further, in one of the first attempts to explain the clique structure of the national interlock network, Allen (1978) speculated that localization of the corporate interlock network is a product of the localization of resource dependence relations in the national economy. He reasoned that corporations tend to compete and transact primarily with other firms headquartered near them and that corporations reduce resource uncertainty by interlocking with local competitors and transaction partners as opposed to firms headquartered in distant locales. This implies that the effects on interlocking of

⁴ Elite theorists argue that large corporations also have more prominent and powerful directors and that it is the prominence and power of their directors—rather than the resource uncertainty and constraint they generate for other firms—that accounts for the greater number of interlocks they maintain (Useem 1979).

competitive uncertainty, interindustry constraint, large size, and industrial diversity should be most evident in connection with local interlocking.

Researchers have found that the level of interlocking within and between manufacturing sectors is directly related to the level of resource dependence—based uncertainty and constraint in and between them (Pfeffer and Salancik 1978; Burt et al. 1980; Burt 1980, 1983). However, in firm-level analyses, no one has observed an association between the number of interlocks linking corporations and the level of dependence between them (Galaskiewicz et al. 1985; Palmer et al. 1986; Mizruchi 1989, 1992). Further, there is conflicting evidence regarding the impact of financial dependence on financial interlocking, especially in connection with corporate reliance on long-term debt. Some researchers find no association between long-term indebtedness and financial interlocking (Allen 1974; Mizruchi and Stearns 1988), one finds a positive association (Pfeffer 1972), and one even reports a negative relationship (Pennings 1980). While it is well established that large firms maintain more interlocks than small ones, no one has examined the relationship between diversification and interlocking.

Bank hegemony theory.—According to bank hegemony theory, the interlock network reflects and facilitates the control of capital flows in the economy (Mintz and Schwartz 1985). In its crudest form, this theory simply reaffirms the basic resource dependence tenet that interlocks are created to reduce uncertainty and constraint associated with intercorporate transactions, adding that financial dependence is the most problematic form of dependence. Industrial firms are particularly dependent on financial institutions because these institutions control the most fungible and highly concentrated resource in the economy—loan capital. In this view, if there is a relationship between resource dependence and interlocking, it should be most apparent in connection with financial dependence and interlocking.

In its more sophisticated form, bank hegemony theory views interlocks as the backbone of a general economic intelligence and decision-making network in which financial institutions are the key actors. Industrial firms interlock with financial institutions to obtain information about and a voice in decisions regarding capital allocation, independent of their particular need for external supplies of capital. Financial institutions interlock with industrial firms in order to obtain information about and access to possible lending and investment opportunities—those presented by the firm, as well as those presented by other corporations to which these industrials may be connected. These imperatives cause the national network to become internally differentiated into regional clusters consisting of locally headquartered corporations and financial institutions, in which the

latter are *peaks*. The regional clusters are in turn tied together by industrial corporations that are central in national capital flows—large, geographically extensive, debt-dependent corporations called *bridges*. As industrial corporations grow in size, geographic scope, and need for capital, they seek out interlocks with national financial institutions outside their headquarters region. Large geographically dispersed firms seek involvement in economywide capital allocation decisions because they operate at the national level. Heavily indebted firms seek information about and access to national capital markets because their capital needs outstrip local reserves. At the same time, financial institutions—especially national financial institutions—become increasingly interested in interlocking with these large, geographically dispersed, debt-dependent firms in order to gain access to information and expertise related to the local economies in which they are situated. Thus bank hegemony theory implies, in contrast to the traditional resource dependence view, that the effect of large size, geographic dispersion, and indebtedness on interlocking should be most evident in connection with nonlocal financial interlocks.

Urban growth machine theory.—We believe that urban growth machine theory, developed to explain community power structures, is consistent with both the traditional resource dependence and bank hegemony theories and has implications for the study of interlocking (Molotch 1976; Logan and Molotch 1987). Cities, in this perspective, are “growth machines”—institutional complexes geared to the generation of economic and population growth. Local business elites and state officials band together in “growth coalitions” that promote public policies intended to stimulate local growth. At the root of each coalition’s structure is a social network linking local economic elites and public officials (Perrucci and Pilisuk 1970). Interlocking directorates among local firms is one means to participate in this network. We think a corporation’s interest in participating in its headquarters city’s growth coalition is a function of its growth dependence on, economic dominance over, and capacity to exit from its headquarters location (Friedland and Palmer 1984; Palmer and Friedland 1987).

Corporations vary in the extent to which their success hinges on the conditions regulating local growth. Growth dependent corporations, such as those maintaining local production facilities, have an incentive to join their local growth coalition (Molotch 1976). Corporations that funnel a large volume of resources into the economies where they are headquartered, relative to the investments of other firms, are dominant economic actors in their communities (Lincoln 1977). Dominant corporations have less need to participate in local growth coalitions in order to influence local state policies because they possess systemic power. Local coalitions champion dominant firms’ interests in order to sustain the local presence

of the resources they control, allowing them to avoid costly direct action (Crenson 1971; Molotch 1979; Friedland 1983). Corporations embedded in geographically expansive financial and nonfinancial markets possess locational flexibility (Bluestone and Harrison 1982; Harrison 1994). In particular, the more locations in which corporations produce, the better positioned they are to shift production outside their headquarters location, either by downsizing and transferring personnel or by closing and relocating entire plants (Pred 1977). The greater a firm's capacity to exit its headquarters locale, the less it needs to participate in the local growth coalition. Like economic dominance, mobility is a source of systemic power that causes local growth coalitions to pursue a corporation's interests. Thus, growth dependence should be associated with high levels of local interlocking, while dominance and mobility should be associated with low levels of local interlocking.

Financial institutions were economically dominant but growth dependent and immobile vis-à-vis their headquarters locales in the 1960s (Friedland and Palmer 1984). Local financial institutions regulated regional growth, channeling the flow of local private and public capital investment. Their prosperity was largely dependent on the success of local nonfinancial firms that generated the deposits and loans upon which their money-making activities hinged. Finally, their operations were geographically circumscribed by law. Hence, financial institutions were at the center of the urban growth machines of the 1960s. Growth dependent, immobile, and nondominant corporations should have been particularly interested in joining with local financial institutions to foster or control local growth because local financials were dominant economic actors. And, financial institutions should have been particularly willing to work with local industrial firms to this end because they were growth dependent and immobile. Thus, the effects of growth dependence, economic dominance, and capacity to exit on local interlocking should have been most evident in connection with ties with financial institutions.

Only one study has explored the relationship between the geography of corporate operations and interlocking—and only tangentially at that. Allen (1974) observed that financial institutions had a higher percentage of local interlocks than did industrial corporations and attributed this to the presumed spatial constriction of financial institution operations.

Ownership Relations

Separation of ownership and control.—The traditional resource dependence perspective assumes that corporate leaders seek stable growth, which leads to the pursuit of environmental certainty, which in turn leads

to interlocking (Pfeffer and Salancik 1978). However, owner managers may be less preoccupied with stable growth and more focused on profit maximization than professional managers. Owner managers accumulate wealth in the form of stock appreciation tied to their firm's profitability, while professional managers earn salaries pegged to their firm's size (Galbraith 1971; Bell 1973). Additionally, the expertise and friendship thesis and the upper-class social interaction view presume that corporate managers seek protection from threats to their tenure, which leads to the pursuit of trusted business allies, which in turn leads to interlocking. However, owner managers may have less need to array outside directors as allies in times of crisis. The tenure of owner managers is more secure than that of salaried professional managers because their stock holdings are a barrier to corporate takeover (Barber, Palmer, and Wallace 1995; Palmer et al. 1995). Thus, corporations whose stock is concentrated in the hands of a few related individuals involved in management should maintain fewer interlocks than firms whose stock is held by unrelated and uninvolved outsiders. Some owner managers reside in their firm's headquarters locale, while others live in distant places. Local owner managers are more involved in their firms' day-to-day affairs, and this provides them with information and contacts that bolster their political security (Pfeffer 1992). This implies that the negative effect of managerial ownership on interlocking should be most evident in connection with ownership by locally resident managers.

Intercorporate ownership.—Researchers have long thought that interlock relations at least partly shadow intercorporate ownership relations (Perlo 1957; Zeitlin 1974; Mintz and Schwartz 1981b; Mizruchi 1982). Corporations and financial institutions can prevail on the firms they own to establish interlock relations with them, in order to monitor these firms' affairs. They may also press the several firms they own to interlock with each other in order to coordinate their sphere of influence (Sweezy 1953). Bank control theorists believe the relationship between financial ownership and interlocking is particularly strong (Fitch and Oppenheimer 1970; but see Kotz 1978). Thus, corporations whose stock is owned by other firms, especially firms that control still other large firms, may maintain more interlocking directorates than firms not embedded in intercorporate stockholding relations. Some believe localized corporate interlock patterns reflect localized intercorporate ownership relations, anchored by local capitalist families. In this view, interlocks arise as the result of ties among locally situated directors associated with the same families, in order to facilitate coordination between the local firms under these families' common control (Sweezy 1953). Thus, ownership by locally headquartered industrial and financial institutions should give rise to local interlocks,

and ownership relations with firms whose headquarters are not local should foster nonlocal interlocking. That is, the predicted positive effect of intercorporate ownership on interlocking should be most evident in the relationship between local intercorporate ownership and local interlocking and in the association between nonlocal intercorporate ownership and nonlocal interlocking.

DATA AND METHOD

Sample and Dependent Variables

Our sample consists of the 500 largest U.S. industrial corporations in 1964, as reported by *Fortune* (1965). We focus on 1964 because data on many difficult-to-collect variables (in particular interlocking directorates, plant locations, and ownership relations) were already available for this year. Five agricultural cooperatives and five wholly owned subsidiaries of parent organizations were excluded from the sample because their financial data were not comparable to those of other top 500 firms. We measured the number of total, local, and nonlocal directional interlocks these corporations maintained with other industrial firms and financial institutions. We focus on directional interlocks, which connect two firms when an inside director (an officer or principal owner) of one also sits on the board of another, because they are considered the most purposeful and the most consequential form of interlocking.⁵ Interlocks linking corporations headquartered in the same city were considered *local*, while those connecting firms headquartered in different cities were considered *nonlocal*.⁶ Infor-

⁵ Directional interlocks are more likely than nondirectional interlocks to be the result of conscious initiatives on the part of connected firms because nondirectional interlocks are automatically created between two firms when they both appoint the same director from a third firm to their boards. Hence, directional interlocks are referred to as "primary," while nondirectional interlocks are termed "secondary" (Mintz and Schwartz 1985, p. 186). Further, many believe that directional interlocks are more likely than nondirectional interlocks to serve as instruments of intercorporate influence and information exchange because they are created by persons who are more willing and able to serve as representatives of one of the firms they connect (Mizruchi and Bunting 1981; Ornstein 1984; Palmer 1983; Palmer et al. 1986; Stearns and Mizruchi 1987; but see Palmer, Jennings and Zhou 1993; Palmer et al. 1995). Regional clusters or cliques within the corporate network are most apparent in analyses restricted to directional interlocks (Mizruchi 1982; Mintz and Schwartz 1985).

⁶ Use of city boundaries, rather than Standard Metropolitan Statistical Area (SMSA) or regional boundaries, to distinguish local from nonlocal interlocks allows us to more precisely test theories about the role interlocks play in organizing business participation in the urban power structure—insofar as SMSAs and regions are not political entities per se. Further, use of SMSAs would have been problematic from an empirical standpoint. Some of our independent measures—in particular, those related to the

mation on corporate interlocks was obtained from a data archive compiled by the MACNET research group at the State University of New York, Stony Brook (Atwood et al. 1985). It contains information on the interlock connections our corporations maintained with other firms that were among the 500 largest U.S. industrial corporations in any year between 1962 and 1973, the 50 largest commercial banks, insurance companies, and diversified financials over the same period, and the 57 leading investment banks in the 1960s. The headquarters locations of the 1,131 MACNET corporations were obtained from the Standard and Poor's *Directory of Corporations, Directors, and Executives* (1965).

Independent Variables

The number of other industrial firms and financial institutions headquartered in a corporation's headquarters city was measured by the number of MACNET organizations of each type headquartered there in 1964. The presence of exclusive upper-class social clubs in a corporation's headquarters city was measured by a dummy variable, coded "1" if at least one of Domhoff's clubs was situated there and "0" otherwise (Domhoff 1970, pp. 233–24).⁷

geography of production—were not readily available at the SMSA level. And for firms not headquartered in an SMSA (of which there were many in 1964), many variables—e.g., those related to the number of other firms headquartered in a corporation's headquarters location—could not have been calculated at all. Finally, use of state and regional boundaries would have been problematic, insofar as both might dampen and even distort measurement of the covariate effects that tap social processes tied to local commute distances—such as the effect that local upper-class clubs may have on interlocking (Kono 1990). Distant cities may be located in the same large states and regions, while proximate cities may be located in different states and regions. For example, if we used state boundaries, interlocks between firms headquartered in New York City and Buffalo (N.Y.) would have been considered local interlocks, while interlocks between firms headquartered in New York City and Trenton (N.J.) would have been considered nonlocal, even though Trenton is much closer than Buffalo to New York City.

⁷ Domhoff (1970) used four methods to identify his 40 exclusive metropolitan upper-class social clubs. Clubs were considered exclusive and upper class in character if they met at least one of the following criteria: (1) membership in them was associated with the likelihood that persons listed in *Who's Who in America* were also listed in one of the metropolitan social registers widely recognized to be indicative of upper-class status, (2) they were identified as exclusive upper-class institutions by newspaper society page editors surveyed in cities that possessed a local chapter of the Junior League, a national upper-class service organization, (3) the officers/partners of the leading corporations, banks, foundations, and law firms in 15 cities not covered in the correlational analysis described above tended to belong to them, and (4) they were identified as upper-class institutions by the authors of one of a small number of community case studies (Baltzell 1958; Kavalier 1960). We focus on exclusive upper-class social clubs rather than more inclusive class heterogeneous clubs because their presence in or absence from a city can be used to differentiate headquarters locations on

We measured competitive uncertainty by the absolute value of the difference between the four-firm concentration ratio in a corporation's primary input-output (IO) sector and the average concentration ratio for all manufacturing IO sectors in 1963. Competitive uncertainty is considered greatest in industries with intermediate concentration levels. In unconcentrated industries, other firms' impact on a corporation's market chances is slight. In highly concentrated industries, other firms' impact on a corporation's market chances are great but their behavior is easily monitored and predicted (Pfeffer 1972; Pfeffer and Salancik 1978). A corporation's principal IO sector was determined by translating its primary four-digit standard industrial classification (SIC) code, obtained from the Center for Research in Stock Price's 1991 NYSE/AMEX Master Stock File, into one of 77 IO categories, using the Bureau of the Budget's *Standard Industrial Classification Manual* for 1967 and 1972. The IO sector concentration scores were taken from Burt (1986b, pp. 15–16).

Buyer-supplier constraint was measured by the extent to which a corporation's primary IO sector transacted with only a few other highly concentrated IO sectors (represented by the "aggregate supplier/consumer constraint" scores computed by Burt [1986b, pp. 16–17]), the concentration level in a firm's primary industry and the interaction of these two variables—all measured in 1963. Buyer-supplier constraint is considered greatest in industries that transact with only a few other sectors that themselves are highly concentrated. In such industries, corporations cannot easily pit prospective transaction partners against each other to obtain favorable terms of exchange. However, total buyer-supplier constraint, as well as the impact of aggregate supplier/consumer constraint on total constraint, is expected to be reduced in industries where concentration is high. Corporations producing in oligopolistic industries enjoy countervailing power over other firms with which they transact. Financial constraint was measured by a firm's debt-to-equity ratio (long-term debt/common equity) and quick ratio (current assets/current liabilities). Financial dependence is greatest when the debt-to-equity ratio is high and the quick ratio is low (Stearns and Mizruchi 1993). Information on long-term debt, common equity, current assets, and current liabilities was obtained from the Standard and Poor's COMPUSTAT annual data file or from Moody's *Industrial Manual* for 1965.

Industrial diversity was measured by the number of two-digit SIC industries in which corporations produced in 1964, as reported in Standard

the basis of the completeness of their upper-class organization. All cities possess numerous social clubs of different types, but only 30 of the 175 cities in which the 1964 top 500 corporations were headquartered possessed one of the clubs on Domhoff's list.

and Poor's *Register of Corporations, Directors, and Executives* (1965). Corporate size was measured by the volume of a firm's total assets in 1964, as reported in the COMPUSTAT annual data file or Moody's *Industrial Manual* (1965).

Three dimensions of corporate geography were tracked to provide information on each firm's geographic complexity and centrality in national capital flows, as well as each firm's growth dependence on, economic dominance over, and capacity to exit from its headquarters city—all for 1964. We calculated the proportion of production facilities corporations situated in their headquarters city. The higher this percentage, the less geographically complex a corporation's operations, the more peripheral its position in national capital flows, and the greater its growth dependence on its headquarters city. We also calculated the percentage of all production facilities situated in a corporation's headquarters city that was owned by the corporation. The higher this percentage, the greater a corporation's dominance over its headquarters city. Lastly, we calculated the number of U.S. states in which each corporation operated production facilities. The higher this number, the more geographically complex a corporation's operations, the more central its position in national capital markets, and the greater its capacity to exit from its headquarters city. Corporate plant locations were obtained from *The Fortune Plant and Product Directory* (Fortune 1963).

Data on managerial stock ownership were obtained from *The Managerial Revolution Reassessed* (Burch 1972), which covers the largest 500 U.S. industrial corporations in 1964. Blocks of stock owned by a firm's management were quantified as a percentage of outstanding common stock (when greater than 3%). Blocks associated with owner managers who lived in the firm's headquarters SMSA were distinguished from those associated with owner managers residing elsewhere. Owner manager residences were primarily obtained from *Who's Who in Commerce and Industry* (Marquis 1965). Data on corporate stock ownership were also taken from Burch (1972). Again, ownership blocks were quantified as a percentage of outstanding common stock. Further, blocks were distinguished on the basis of whether or not the owning corporations were headquartered in the owned firm's headquarters city. We also constructed a composite measure of the number of other corporate ownership relationships that corporations maintained with other industrial firms—which included the number of other firms in which they owned a recognizable share and the number of other firms owned by the same family or firm that owned a recognizable share in them. These "other corporate ownership relationships" were also distinguished on the basis of whether or not the related corporations were headquartered in the same or a different city as the focal corporation. Data on financial ownership were obtained from *Com-*

mercial Banks and Their Trust Activities (U.S. Congress 1968), which covers the largest 500 U.S. industrial corporations in 1966. Each block of stock (with at least partial voting rights) owned in a corporation by a financial institution was quantified as a percentage of the firm's total outstanding common and preferred stock. Further, ownership blocks were distinguished on the basis of whether the owning financials were headquartered in the owned firm's headquarters city. Missing financial ownership data was the primary reason for case exclusion—45 of the 63 firms eliminated from the financial interlock analyses were missing bank ownership information.⁸

The independent variables measured for this study, as well as the theoretical constructs they index, the theories from which these constructs are drawn, and the variables' predicted effects on interlocking are all depicted in table 2. Some of our variables index more than one theoretical construct drawn from different theories. In these cases, we offer more than one prediction regarding a variable's effect on interlocking. For example, the number of U.S. states in which a corporation operates production facilities is one indicator of the complexity of its geographical structure. As such, traditional resource dependence theorists would predict it to be positively related to total interlocking (local and nonlocal alike). However, geographically dispersed production also indexes a firm's centrality in national capital flows. As such, bank hegemony theorists would expect the positive association between it and interlocking to be particularly evident in connection with nonlocal interlocking. Finally, geographically dispersed production also indexes a firm's capacity to exit its headquarters locale. As such, growth machine theorists would expect it to be negatively associated with local interlocking. Tables 2 and 3 provide descriptive statistics for and correlations among the independent variables described above.⁹

⁸ Because bank stock ownership data were missing primarily for firms that had dropped out of the Fortune 500 between 1964 and 1966 due to acquisition or declining asset rank, the sample used for our financial interlock analyses is probably slightly biased in favor of the largest and best performing of the 1964 Fortune 500 firms. This possible bias, though, had no apparent influence on the results we obtained. We conducted analyses parallel to those reported here that omitted the financial ownership variables and thus reduced the level of missing data. The results of these parallel analyses were essentially identical to those we report here. We did not attempt to measure other corporate ownership relationships that might arise from financial ownership (e.g., common ownership by the same financial institution) because doing so would require considerable effort and our results revealed little evidence that direct financial ownership influenced interlocking.

⁹ We also tested for the possible effects of several other variables on interlocking: use of the multidivisional form, ownership by locally and nonlocally resident persons not represented in management, and ownership by locally and nonlocally headquartered

Method

The numerical distributions of our dependent variables (e.g., local industrial, nonlocal industrial, local financial, etc.) are highly skewed, with most firms maintaining no interlocks of the type in question, many firms maintaining a few such interlocks, and a handful of firms maintaining large numbers of interlocks. Preliminary analyses using Poisson and negative binomial count regression models (Cameron and Trividi 1986, 1990) indicated that our results were heavily influenced by firms possessing large numbers of interlocks, especially those with extremely high numbers of interlocks. This led us to strongly suspect that the effects of some of our covariates on interlocking were nonmonotonic and, more important, that inclusion in our analyses of firms with extremely high numbers of interlocks—the “biggest linkers”—distorted our count regression results.

To accommodate these problems while retaining the biggest linkers in our study, we used the univariate distributions reported in figure 1 to categorize our firms into three approximately equal sized groups for the analysis of each type of interlock: firms with no interlocks of the type in question, firms with one interlock, and firms with two or more interlocks. The effects of covariates on the likelihood that corporations maintained one rather than no interlocks (the none/one contrast) and two or more rather than no interlocks (the none/two-or-more contrast) were estimated using multinomial logistic regression (Greene 1993; Whitten and Palmer 1996). This allows us to determine whether the covariates measured here were associated with the maintenance of interlocks of particular types, while limiting the potential result-distorting influence of firms that maintain extremely large numbers of interlocks to the none/two-or-more contrast. If covariate effects are apparent in connection with both the none/one and none/two-or-more contrasts, we infer that they influence interlocking of the type in question—serving to distinguish firms that maintain no interlocks of that type from those that maintain one or more. If covariate effects are apparent in connection with only one of the two contrasts, we consider the results less conclusive. If covariate effects are apparent only in connection with the none/two-or-more contrast, we infer that they might influence interlocking of that type but only at higher levels of interlocking—distinguishing firms that maintain no interlocks from those that maintain two or more interlocks. And if covariate effects are apparent only in connection with the none/one contrast, we infer that they might influence interlocking of the type in question but that the unique attributes

firms that were not in the MACNET data set, most of which were foreign. None of these variables had noticeable effects on total, local, or nonlocal interlocking with other industrial firms or financial institutions.

TABLE 2
HYPOTHEZED EFFECTS OF COVARIATES ON INTERLOCKING OF LARGE U.S. INDUSTRIAL CORPORATIONS, 1964

Variable*	Construct†	Mean	SD	Total Interlocks	Local Interlocks	Nonlocal Interlocks
No. of other industrial corporations headquartered in firm's headquarters city (NIC)	Proximity to other corporate headquarters (expertise and friendship)	54.795	72.312	+	++	+
No. of financial institutions headquartered in firm's headquarters city (NFC)	Proximity to financial institution headquarters (expertise and friendship)	20.480	28.698	+	++	+
Presence of an exclusive upper-class club in firm's headquarters city (CLUBC)	Proximity to exclusive upper-class clubs (upper-class social interaction)	618	486	+	++	+
Mean-deviated concentration level in firm's primary industry (ADCONC)	Decreasing competitive uncertainty (traditional resource dependence)	.126	.104	-	--	-
Industry concentration level in firm's primary industry (ICC)	Oligopoly power over firms in other industries (traditional resource dependence)	.406	163	-	--	-
Industry constraint level in firm's primary industry (ICS)	Interindustry constraint posed by firms in other sectors (traditional resource dependence)	.050	.036	+	++	+
ICC × ICS	Interaction of oligopoly power and interindustry constraint (traditional resource dependence)	-.001	.007	-	--	-
Debt-to-equity ratio (DEQ)	Dependence on long-term debt (traditional resource dependence)	.310	471	+	++	+
	Central position in capital flows (bank hegemony)			+	+	++

Quick ratio (QUICK)	Independence from short-term debt (traditional resource dependence)	2 955	1 011	-	--	-
	Peripheral position in capital flows (bank hegemony)			-	-	--
Total assets, in \$100,000 (ASSETS)	Uncertainty generated for other firms (traditional resource dependence)	4 593	10 250	+	++	+
	Central position in capital flows (bank hegemony)			+	+	++
No. of two-digit SIC industries in which firm produces (NSIC2)	Industrial diversity of product mix (traditional resource dependence)	4 029	2 813	+/-	++/-	+/-
% firm's plants situated in its headquarters city (PCPHQC)	Simplicity of geography of corporate production (traditional resource dependence)	.132	.207	-	-	-
	Peripheral position in capital flows (bank hegemony)			-	-	--
	Growth dependence on headquarters city (growth machine)			0	+	0
No. of states in which firm owns plants (NSTP)	Complexity of geography of corporate production (traditional resource dependence)	9 407	7 572	+	+	+
	Central position in capital flows (bank hegemony)			+	+	++
% plants in headquarters city owned by firm (PTPHQC)	Capacity to exit from headquarters city (growth machine)			0	-	0
	Economic dominance in headquarters city (growth machine)	098	.226	0	-	0
% firm's stock owned by a locally resident group of individuals represented in management (PSLROM)	Local insider ownership (separation of ownership and control)	9 246	18 109	--	--	--

TABLE 2 (Continued)

Variable*	Construct†	Mean	SD	Total Interlocks	Local Interlocks	Nonlocal Interlocks
% firm's stock owned by a nonlocally resident group of individuals represented in management (PSNLROM)	Nonlocal insider ownership (separation of ownership and control)	1 897	7 933	-	-	-
% firm's stock owned by locally headquartered industrial corporations (PSLHQC)	Local corporate ownership (intercorporate and financial control)	853	7 065	+	++	+
% firm's stock owned by nonlocally headquartered industrial corporations (PSNLHQC)	Nonlocal corporate ownership (intercorporate and financial control)	713	6 098	+	+	++
No of other locally headquartered firms to which corporation is linked via ownership relations (ORLHQC)	Other local corporate ownership relationships (intercorporate and financial control)	065	346	+	++	+
No of other nonlocally headquartered firms to which corporation is linked via ownership relations (ORNLHQC)	Other nonlocal corporate ownership relationships (intercorporate and financial control)	.147	.597	+	+	++
% firm's stock owned by locally headquartered financial institutions (PSLHQF)	Local financial institution ownership (intercorporate and financial control)	.743	2 573	+	++	+
% firm's stock owned by nonlocally headquartered financial institutions (PSNLHQC)	Nonlocal financial institution ownership (intercorporate and financial control)	1 326	4 314	+	+	++

NOTE.—The symbols for predicted direction and magnitude of effects are + = positive effect, ++ = particularly evident positive effect, - = negative effect, --- = particularly evident negative effect, 0 = no prediction

* The letters in parentheses are the variable abbreviations

† The related theory appears in the parentheses

TABLE 3

CORRELATIONS

	NIC	CLUBC	PCPHQC	PTPHQC	NSTP	PSNLROM	PSLROM	PSNLHQC	PSLHQC	ORNLHQC	ORLHQC	PSLHQC	PSNLHQC	ADCONC	ICC	ICS	IOCKICS	DEQ	QUICK	NEICA	ASSISTS
NFC
NIC		55*	-23*	-28*	18*	-05	-09*	-03	01	-05	09	11*	-17*	00	-06	03	-06	04	11*	07	10*
CLUBC	-		58*	-31*	18	-05	-10*	-03	02	-04	10*	13*	-17*	.00	-08	02	-06	04	12*	07	10*
PCPHQC	-			-24*	20*	-04	-10*	.00	02	.01	10*	19*	-18*	.05	-04	03	-06	.06	00	.11*	08*
PTPHQC	-				30*	03	.13*	-05	-02	-05	-11*	-01	.06	.09*	.06	-07	.05	-07	.02	-20*	-11*
NSTP	-					-01	09	-03	-05	.00	-06	-11	08	00	06	-05	07	-07	01	08	05
PSNLROM	-					-03	-13*	-06	-04	04	.06	00	-09	-10*	-06	-18*	-10	03	01	36*	19*
PSLROM	-					-12*	-07	07	06	-02	-02	03	09	02	02	02	01	-03	-03	02	-05
PSNLHQC	-					06	-07	-01	.04	-06	-06	.01	.05	-07	-05	-04	-04	-01	.03	-11*	-06
PSLHQC	-					-01	-10	20*	03	-03	03	-03	-03	.00	02	.15*	-06	.01	03	-06	01
ORNLHQC	-					16*	23*	-01	01	03	01	01	01	03	01	01	.04	13*	-09*	04	-03
ORLHQC	-					17*	03	-01	03	-04	.06	-07	02	.05	03	.02	-07	02	03	-03	14*
PSLHQC	-					27*	03	-03	-03	-03	03	03	-03	-03	03	02	.05	-08	07	05	07
PSNLHQC	-					-06	-06	-06	00	03	06	-07	02	00	03	06	-07	07	-04	03	03
ADCONC	-					-01	-01	-01	-01	-01	-01	-01	-01	-01	-01	-01	-01	-04	-06	-07	-04
ICC	-					50*	50*	50*	50*	50*	50*	50*	50*	50*	50*	50*	50*	50*	50*	50*	50*
ICS	-					17*	17*	17*	17*	17*	17*	17*	17*	17*	17*	17*	17*	17*	17*	17*	17*
IOCKICS	-					08	08	08	08	08	08	08	08	08	08	08	08	08	08	08	08
DEQ	-					-13*	-13*	-13*	-13*	-13*	-13*	-13*	-13*	-13*	-13*	-13*	-13*	-13*	-13*	-13*	-13*
QUICK	-					05	05	05	05	05	05	05	05	05	05	05	05	05	05	05	05
NEICA	-					00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00

* $P < .05$ (two-tailed test)

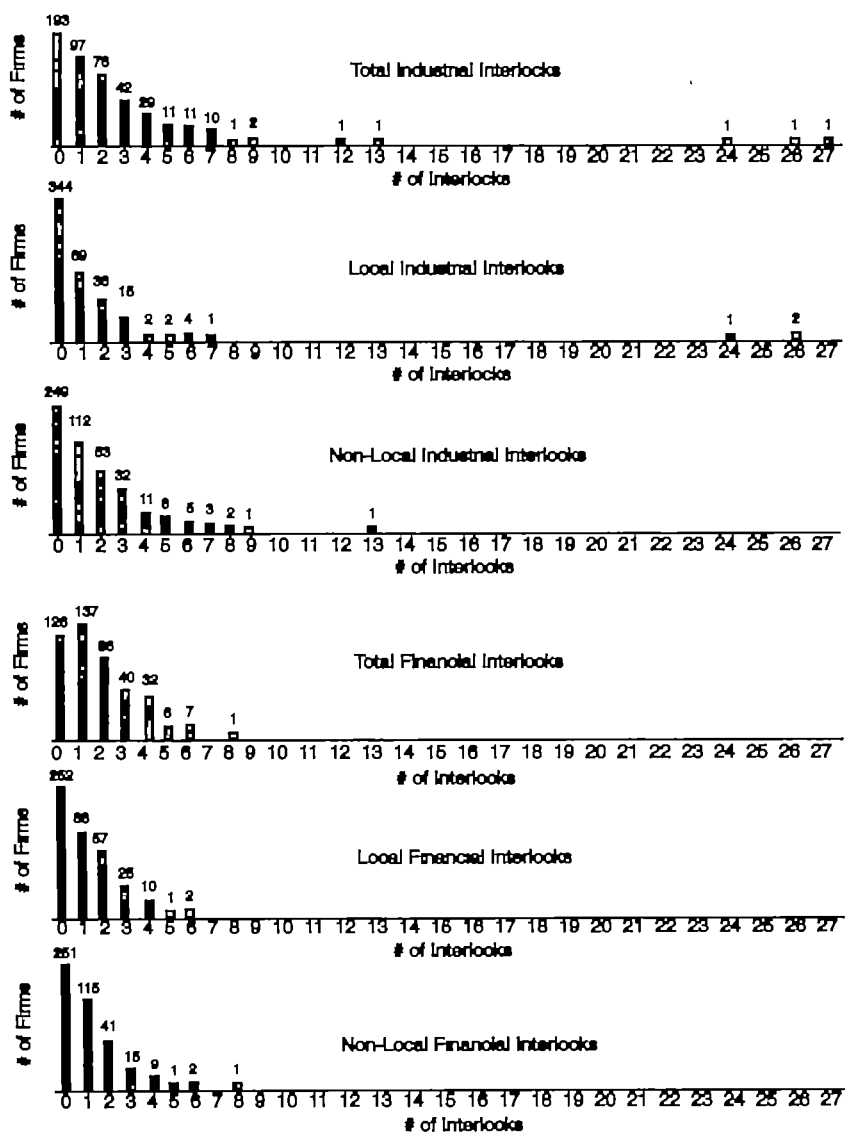


FIG. 1.—Frequencies for total, local, nonlocal industrial and financial interlocks.

of the biggest linkers masks this effect among firms with more than one interlock.¹⁰

This approach to analyzing our data is empirically rather than theoretically motivated. Further, it forfeits the opportunity to estimate how our covariates influence variation in interlocking among firms that maintain two or more interlocks. However, extensive experimentation with alternative approaches to analyzing our data suggest that this method produced the most straightforward and faithful representation of the relationships between interlocking and its hypothesized determinants, as they are conceptualized in this study.¹¹

RESULTS

Comparison of Total with Local and Nonlocal Interlocking

The fundamental assumptions motivating this study are that interlocking is a spatial phenomenon and that previous research, because it fails to take this into account, may have overlooked or misconstrued determinants of interlocking. To determine the extent to which this may be the case, we conducted an aspatial analysis of total industrial and financial interlocking, presented in table 4. The results of this analysis can be compared with the results of our separate analyses of local and nonlocal interlocking, presented in tables 5 and 6. If the aspatial analysis of interlocking misapprehends the effects of hypothesized determinants of interlocking, the results of the separate local and nonlocal interlock analyses should differ substantially from the results of the total interlock analysis. Most

¹⁰ The effects of covariates on the likelihood that firms maintained two or more rather than one interlock of a particular type can be obtained by subtracting the coefficient estimates for the none/two-or-more contrast from the estimates for the none/one contrast. We chose not to report these effects because they were seldom statistically significant, indicating that the factors regulating interlocking have a greater influence on whether or not corporations establish a first interlock of a particular type than they do on whether or not firms establish a second one.

¹¹ We originally estimated Poisson and negative binomial count regression models with and without the biggest linkers (e.g., in the local industrial interlock analysis, firms with 24 or more interlocks) and other big linkers (e.g., in the local industrial interlock analysis, those with four or more interlocks) and identified those covariate effects that appeared to be stable regardless of whether or not these biggest or big linkers were included in the analysis. This method was abandoned because it entailed repeated reconfiguration of our sample based on the values firms exhibited on the dependent variables, producing results that were cumbersome to report. We also estimated binomial logit models run on the full sample. This method was abandoned because it did not take into account the possible results-distorting effects of the biggest linkers. Although we deemed these alternative approaches inferior to the method we describe above, we found it reassuring that the results of these alternative analyses generally agreed with those we report in the text.

TABLE 4
TOTAL INTERLOCKING WITH OTHER INDUSTRIAL CORPORATIONS AND FINANCIAL INSTITUTIONS

VARIABLE	INDUSTRIAL INTERLOCKING		FINANCIAL INTERLOCKING	
	ln(One/None)	ln(Two or More/None)	ln(One/None)	ln(Two or More/None)
No. of other industrial firms headquartered in firm's headquarters city	0039*	.0015	-.0019	-.0090
Presence of an exclusive upper-class club in firm's headquarters city	-.4315	-.0572	.5384*	1.2377**
Mean-deviated concentration level in firm's primary industry	-1.1324	-1.6228		
Industry concentration level in firm's primary industry (ICC)	6057	1.0199		
Industry constraint level in firm's primary industry (ICS)	-6753	2.7457		
ICC x ICS	-38,4050**	-22.841		
Debt-to-equity ratio8726**	3118
Quick ratio			-.0540	-.0602
Total assets (in \$100,000)	-.0258	.0315*	0461	1053**
No. of two-digit SIC industries in which firm produces1566**	.1857**	1036**	0290
% firm's plants situated in its headquarters city	4189	4032	8258	-.0266
No. of states in which firm owns plants	-.0012	0179	.0225	.0597**
% plants in headquarters city owned by firm5007	-.5532	.7565	8093
% firm's stock owned by a locally resident group of individuals represented in management	-.0180**	-.0167**	0012	0041

% firm's stock owned by a nonlocally resident group of individuals represented in management	-.0370**	-.0188*	.0292	.0069
% firm's stock owned by locally headquartered industrial corporations	.0349*	.0067		
% firm's stock owned by nonlocally headquartered industrial corporations	.1059*	.0855		
No. of other locally headquartered firms to which corporation is linked via ownership relations	-.8686	1.2876**		
No. of other nonlocally headquartered firms to which corporation is linked via ownership relations	-.5048	.2397		
% firm's stock owned by locally headquartered financial institutions			-.0156	.0326
% firm's stock owned by nonlocally headquartered financial institutions			.0467*	.0471*
Constant	-1.1266	-1.4154	-1.3207	-1.4651
LR test statistic (df)		102.83 (42)**		81.189 (42)**
% correctly predicted		50.9		50.1
% reduction in error		17.6		17.5
N		477		435

* $P < .10$ (one-tailed tests for unidirectional hypotheses, two-tailed test otherwise)

** $P < .05$

TABLE 5

LOCAL AND NONLOCAL INTERLOCKING WITH INDUSTRIAL CORPORATIONS

VARIABLE	LOCAL INTERLOCKES		NONLOCAL INTERLOCKES	
	ln(One/None)	ln(Two or More/None)	ln(One/None)	ln(Two or More/None)
No. of other industrial firms headquartered in firm's headquarters city0009	.0030*	-.0004	.0001
Presence of an exclusive upper-class club in firm's headquarters city . . .	1.8080**	2.1081**	.0276	-.9806**
Mean-deviated concentration level in firm's primary industry	-.5291	.3940	-3.9353**	-2.7511**
Industry concentration level in firm's primary industry (ICC)3891	-.2775	1.9027**	.9730
Industry constraint level in firm's primary industry (ICS)	1.8118	-.1746	3.3772	3.7399
ICC x ICS	-16.9160	-4.9157	-21.9230	-8.5984
Total assets (in \$100,000)	-.0090	.0006	-.0080	.0542**
No. of two-digit SIC industries in which firm produces	1213**	.1346**	.0988	.2207**
% firm's plants situated in its headquarters city1931	.5046	.8617	.5654
No. of states in which firm owns plants0182	.0099	.0331**	.0035
% plants in headquarters city owned by firm	-6.0596*	-19.7460**	.8419	-.2376
% firm's stock owned by a locally resident group of individuals represented in management	-.0223**	-.0089	-.0222**	-.0173**
% firm's stock owned by a nonlocally resident group of individuals represented in management	-.0008	-.0460*	-.0152	-.0097
% firm's stock owned by locally headquartered industrial corporations . .	-.4135	.0413**	.0249	-.0134
% firm's stock owned by nonlocally headquartered industrial corporations	-1.1557	-.0056	.1176*	.1103*
No. of other locally headquartered firms to which corporation is linked via ownership relations0301	1.0350**	-.3152	.1637
No. of other nonlocally headquartered firms to which corporation is linked via ownership relations0106	-.2368	-.0871	.3673**
Constant	-3.5151	-3.6954	-1.9543	-1.6541
L.R. test statistic (χ^2)	162.07 (34)**		108.93 (34)**	
% correctly predicted	72.5		59.3	
% reduction in error	15		14.9	
N	477		477	

* $P < .10$ (one-tailed tests for unidirectional hypotheses, two-tailed test otherwise)** $P < .05$.

TABLE 6

LOCAL AND NONLOCAL INTERLOCKING WITH FINANCIAL INSTITUTIONS

VARIABLE	LOCAL INTERLOCKING		NONLOCAL INTERLOCKING	
	ln(One/None)	ln(Two or More/None)	ln(One/None)	ln(Two or More/None)
No. of other financial institutions headquartered in firm's headquarters city	0083*	0021	-0155**	-0302**
Presence of an exclusive upper-class club in firm's headquarters city	3.3568**	4.5176**	-.3750	-.9497**
Debt-to-equity ratio	-.3599	-1.3071**	.6914**	.8289**
Quick ratio	.0667	.0600	-.0765	-.0705
Total assets (in \$100,000)	.0871**	.0967**	.0288**	.0448**
No. of two-digit SIC industries in which firm produces	.0937	.0230	-.0504	-.0034
% firm's plants situated in its headquarters city	1.9801**	1.7667**	-.3273	-1.1372
No of states in which firm owns plants	.0239	.0587**	.0276*	.0479**
% plants in headquarters city owned by firm	-8.5001**	-5.9030*	1.0682*	1.0930
% firm's stock owned by a locally resident group of individuals represented in management	.0001	-.0012	-.0043	.0016
% firm's stock owned by a nonlocally resident group of individuals represented in management	.0801**	.0661**	.0006	-.0306
% firm's stock owned by locally headquartered financial institutions	.0047	.0500	-.0662	-.1553
% firm's stock owned by nonlocally headquartered financial institutions	.0515	.0562	.0327	.0260
Constant	-4.9755	-5.7026	-.4323	-1.0018
LR test statistic (df)		255.77 (26)**		96.10 (26)**
% correctly predicted		67.4		61.4
% reduction in error		22.4		8.7
N		435		435

* $P < .10$ (one-tailed test for unidirectional hypotheses, two-tailed test otherwise)** $P < .05$

important, we should find that some covariates have no effect on total interlocking but have noticeable effects on either local or nonlocal interlocking. We might even find that some hypothesized determinants influence local and nonlocal interlocking in different ways. We should, however, find few instances of covariates influencing total interlocking but not local or nonlocal interlocking. This is precisely what we observe.

The presence of exclusive upper-class social clubs in a corporation's headquarters city, the proportion of locally situated production facilities that a corporation owns, the number of states in which a corporation produces, being related by ownership to nonlocally headquartered firms, and the mean-deviated concentration level in a corporation's primary industry all influence either local or nonlocal industrial interlocking but not total industrial interlocking. The number of financial institutions headquartered in a firm's headquarters city, the percentage of a corporation's production facilities located in its headquarters city, the percentage of local plants owned by a corporation, and the percentage of a corporation's stock owned by nonresident outsiders all influence either local or nonlocal financial interlocking but not total financial interlocking. Indeed, the presence of upper-class clubs in a corporation's headquarters city has an opposite effect on local and nonlocal industrial interlocking. And the number of financial institutions headquartered in a corporation's headquarters city and the percentage of local plants owned by a corporation have opposite effects on local and nonlocal financial interlocking. Only one covariate effect observed in the total interlock analysis is not also observed in either the local or nonlocal interlock analysis. And this effect, a positive association between ownership by nonlocal banks and total financial interlocking, is only marginally significant (at the .10 level). Thus, analyzing interlocking as an aspatial phenomenon appears to overlook and mischaracterize determinants of interlocking. Next we analyze interlocking as a spatial phenomenon.

Local and Nonlocal Interlocking

Expertise and friendship.—The results provide only weak support for the expertise and friendship explanation of interlocking. Corporations headquartered in cities with many other industrial headquarters were more likely to maintain *local* industrial interlocks and those headquartered in cities with many financial headquarters were more likely to maintain *local* financial interlocks. However, the former effect was only apparent in connection with the none/two-or-more contrast, the latter only in connection with the none/one contrast, and both effects were only marginally statistically significant (at the .10 level). Further, corporations headquartered in cities with many financial headquarters were *less* likely to

maintain *nonlocal* financial interlocks. Thus, being headquartered near large numbers of other industrial and financial headquarters, with easy access to a large supply of qualified and presumably familiar corporate board candidates, may increase a corporation's propensity to interlock with other local industrial firms and financial institutions. However, being headquartered in a major financial headquarters complex decreases a corporation's need to invite the inside directors of nonlocal financial institutions to serve on its board.

Upper-class social interaction.—The results were more consistent with the upper-class social interaction account of interlocking. Corporations headquartered in cities with exclusive upper-class clubs were more likely to maintain local industrial and *local* financial interlocks. However, corporations headquartered in cities with upper-class clubs were *less* likely to maintain *nonlocal* industrial and *nonlocal* financial interlocks. These latter effects were only apparent in connection with the none/two-or-more contrast, suggesting that being headquartered in a city with upper-class clubs only distinguished firms that maintained no *nonlocal* interlocks from those that maintained a substantial number. Thus, exclusive upper-class social clubs appear to provide settings in which managers and directors become intimately acquainted with one another, making local elites particularly attractive candidates for a corporation's board. However, in doing so, local upper-class clubs appear to reduce a corporation's need to recruit for its board the inside directors of nonlocal financial institutions.

Resource dependence relations.—The results indicate that resource dependence relations influence interlocking. Competitive uncertainty, one type of interindustry constraint, and large size increase the likelihood of interlocking. Further, the effect of industrial diversity is evident with respect to both local and nonlocal interlocking. Corporations that produce in many different industries are more likely to maintain both local and *nonlocal* industrial interlocks, although the latter effect is only apparent in the none/two-or-more contrast. This suggests that producing in a large number of industries increases a firm's need to obtain information and influence over its environment and that this need motivates interlocking.

The pattern of these resource dependence effects, though, conforms more closely to bank hegemony theory than the traditional resource dependence perspective. Traditional resource dependence theorists consider financial and nonfinancial dependence to be equally problematic. However we find that the effect of interindustry constraint is only evident in connection with financial dependence. Allen (1978) speculated that regionalized interlocking is a consequence of regionalized resource dependence. However, we find that the effect of financial dependence, as well as the effect of competitive uncertainty, was only evident in connection with nonlocal interlocking. Corporations with high debt-to-equity ratios, al-

though not low quick ratios, were more likely to maintain *nonlocal* financial interlocks. Corporations whose primary industry concentration levels departed from (were lower or higher than) the manufacturing sector average were more likely to maintain *nonlocal* industrial interlocks. Our results are more consistent with the bank hegemony assumption that capital dependence is the most problematic type of dependence and, thus, the kind most likely to stimulate interlocking. They also fit the bank hegemony view that financial dependence underpins corporate centrality in national capital flows, thus increasing the likelihood that industrial corporations serve as bridges between their headquarters locale and the national interlock network.

Moreover, large size was primarily associated with financial, as opposed to nonfinancial, interlocking. Large corporations were more likely to maintain local and nonlocal financial interlocks. Large firms were also more likely to maintain more nonlocal industrial interlocks, but this relationship was only apparent in connection with the none/two-or-more contrast. The traditional resource dependence perspective assumes that larger corporations generate more uncertainty for other firms and thus are more frequently the target of other firms' attempts to acquire information and influence. But the evidence is weak that industrial firms interlock more with large corporations than they do with smaller corporations. The pattern of effects is more consistent with the bank hegemony view that large corporations serve as bridges between other firms in their headquarters locale and the national network by interlocking with nonlocally headquartered financial institutions. If large corporations serve as bridges to national financial institutions, though, they do so by establishing interlocks with local financial institutions as well.

Our results also indicate that the effects of nonfinancial dependence (mean-deviated primary industry concentration level and industrial diversity) were largely restricted to industrial interlocking, while the effects of financial dependence (debt-to-equity ratio) were restricted to financial interlocking. In fact, because preliminary analyses revealed no effect of competitive uncertainty and buyer-supplier constraint on financial interlocking nor any effect of financial dependence on industrial interlocking, indicators of competitive uncertainty and buyer-supplier constraint were dropped from the financial interlocking analyses and indicators of financial constraint were dropped from the industrial interlocking analyses. Resource dependence-based motivations for establishing industrial and financial interlocks appear to be relatively distinct.

Further, two anomalous findings were observed. First, corporations with high debt-to-equity ratios were less likely to maintain local interlocks, although this relationship was only apparent in connection with the none/two-or-more contrast. Pennings (1980) observed a negative relation-

ship between debt ratios and total financial interlocking. He speculated that high debt ratios are indicative of poor performance and that the directors of poorly performing firms find it difficult to obtain seats on the boards of financial institutions. If correct, our results suggest that this dynamic is a local phenomenon. Second, as observed in previous studies, none of the three components of buyer-supplier constraint influenced either local or nonlocal interlocking in the predicted direction. In fact, contrary to theory, corporations producing in industries with high concentration levels were more likely to maintain *nonlocal* industrial interlocks—although this relationship was only apparent in connection with the none/one contrast. We speculate about the meaning of this null result in the conclusion.

The results pertaining to the geography of corporate production provide support for both the bank hegemony and growth machine perspectives. Corporations concentrating a high percentage of their plants in their headquarters city were more likely to maintain *local* financial interlocks. This is precisely what our version of growth machine theory predicted. Firms concentrating their productive activity in their headquarters locale depend on the conditions shaping profitability there and thus have an incentive to participate in the local growth coalition, especially by establishing interlocks with the most central local economic actors—financial institutions.

Firms owning a large proportion of the corporate production facilities in their headquarters city were less likely to maintain *local* industrial and *local* financial interlocks. While these effects were apparent in both the none/one and none/two-or-more contrasts, the effect on industrial interlocking was only marginally significant in the none/one contrast and the effect on financial interlocking was only marginally significant in the none/two-or-more contrast. Because firms owning a large proportion of the plants in their headquarters city economically dominate their headquarters locale, they can eschew direct involvement in the local growth coalition. Corporations that own a large proportion of the plants situated in their headquarters locale may even maintain more *nonlocal* financial interlocks, indicating that local dominance frees them to participate in nonlocal coalitions in which they have a stake. We proffer this interpretation tentatively, though, because the effect of local dominance on *nonlocal* financial interlocking is only apparent in connection with the none/one contrast and is only marginally significant at that.

Finally, there is evidence that corporations producing in a large number of U.S. states were more likely to maintain *nonlocal* industrial and *nonlocal* financial interlocks and possibly *local* industrial interlocks. However, the effect of geographic dispersion on nonlocal financial interlocking is only marginally significant in the none/one contrast. Further, the effect

on nonlocal industrial interlocking is only evident in the none/one contrast and the effect on local financial interlocking is only evident in the none/two-or-more contrast. This pattern is most consistent with the bank hegemony view, which argues that firms with geographically extensive production structures are central to national capital flows and thus are likely to serve as bridges between their headquarters locale and the national interlock network, primarily by interlocking with national financial institutions.

Ownership relations.—The effects of managerial ownership on industrial interlocking correspond to our expectations. Ownership by managers residing in the headquarters locale reduced the likelihood that corporations maintained local and *nonlocal* industrial interlocks, although the effect on local interlocking was only apparent in the none/one contrast. Nonlocal insider ownership also reduced the likelihood that firms maintained *local* industrial interlocks; however, this effect was only apparent in the none/two-or-more contrast and was only marginally significant. This suggests that owner managers have less incentive to reduce uncertainty plaguing their firms and/or less need to protect themselves against challenges threatening their tenure. It also supports our speculation that locally resident owner managers are more involved in the day-to-day affairs of their firms and are thus more politically secure in their positions. This diminished concern with uncertainty reduction, coupled with relative political security, makes locally resident insiders less interested in establishing interlock relationships between their firm and others.

The effects of managerial ownership on financial interlocking, though, diverge from our expectations. Ownership by locally resident managers did not reduce either local or nonlocal financial interlocking. Perhaps locally resident owner managers are as motivated as their professional counterparts to establish relationships to financial institutions because the control of capital flows is a particularly pressing imperative. In fact, ownership by nonlocal owner managers increased the likelihood that corporations maintained *local* financial interlocks. Nonlocal residence may preclude close and frequent contact with local financial elites and thus reduce access to information about and influence over local capital flows. If so, nonlocal owner managers may experience a heightened need to link their firms with local financial institutions, in order to compensate for their lack of participation in the local social networks in which financial elites are assumed to be central.

The effects of intercorporate ownership on industrial interlocking are largely consistent with our expectations. Being owned by locally headquartered industrial firms and being otherwise related by ownership to locally headquartered industrials (e.g., being owned by the same family or firm as another local firm) increased the likelihood that corporations

maintained *local* industrial interlocks. Being owned by nonlocally headquartered industrial firms and being otherwise related by ownership to nonlocally headquartered industrials increased the likelihood that firms maintained *nonlocal* industrial interlocks. However, almost all of these effects were apparent only in connection with the none/two-or-more contrast. While corporations tend to interlock with firms headquartered in locations containing firms to which they are related by ownership, the influence of intercorporate ownership apparently operates only at the upper end of the interlock range—serving to distinguish corporations maintaining no interlocks from those maintaining two or more. Bank ownership, though, has no impact on financial interlocking. This is consistent with past research, if not bank control theory, suggesting that bank ownership and financial interlocking are only very loosely linked (Kotz 1978). Perhaps banks possess other mechanisms to monitor and control the firms they own and to which they may tend to lend capital (e.g., legal devices such as receiverships).

DISCUSSION

Interlocking as a Spatial Phenomenon

Interlocking is a spatial phenomenon. The spatial structure of corporate relations—the geography of corporate headquarters location, plant sites, and ownership relationships—is a critical determinant of interlocking. Further, interlocks are spatially structured. To understand the determinants of interlocking, one must take that spatiality into account. While a few hypothesized determinants influenced both local and nonlocal interlocking in the same way, most had significant effects on local *or* nonlocal interlocking but not both. In some instances, there was no discernible relationship between a hypothesized determinant and *total* interlocking, while there were significant relationships between the factor and local *or* nonlocal interlocking. In several cases, covariate effects ran in opposite directions. In a few instances, opposing covariate effects on local and nonlocal interlocking counterbalanced one another. In these cases, had we not distinguished between *local* interlocking and *nonlocal* interlocking, we would have erroneously concluded that the hypothesized determinants had no influence on interlocking at all.

Thus, previous models of interlocking have been misspecified. When researchers study interlocking without regard to the geographic structure of firms and the relationships in which they are embedded, they ignore important determinants of interlocking. And when researchers study interlocks without regard to the location of the corporations they connect—failing to distinguish ties linking firms whose directors work and live proximate to one another (who likely are involved in dense, informal, and

inexpensive networks of social interaction) from those linking firms whose directors work and live distant from one another (who likely communicate and interact infrequently, in non-face-to-face settings, at greater cost)—they confound two different phenomenon, with different determinants and possibly different functions.

Understanding local interlocking.—Interlocks are concentrated among firms headquartered in the same locale. What are the processes that underpin local interlocking? Local interlocking, as distinct from nonlocal interlocking, is most clearly a function of the presence of exclusive upper-class clubs in a firm's headquarters city. Proximity to the headquarters of other large industrial firms and financial institutions may allow local corporate directors to become acquainted with one another. However, local upper-class clubs facilitate intense and intimate local elite interaction through which directors develop trust and thus lay the foundation for local interlocking. It is not sufficient for elites to be proximate; they must have an institutional context conducive to interaction if they are to select one another to serve on each others' boards. Many researchers have argued that the corporate elite is organized around upper-class social institutions. However, they have focused on how upper-class club affiliations, conceptualized as personal attributes, affect the likelihood that individual elites are central within the capitalist class, as reflected in the acquisition of corporate top management positions, corporate board seats, and public policy-making group memberships (Soref 1976, 1980; Useem 1979; Useem and Karabel 1986; Bearden and Mintz 1987; Johnson and Mintz 1989). Our results suggest that upper-class social institutions are important ecological facts and that proximity to them shapes corporate interlocking.

What is the broader significance of the local interlocks facilitated by local upper-class clubs? We cannot answer this question definitively, insofar as we have no information on the corporate behaviors or outcomes that might follow from upper-class-club-facilitated local interlocks. However, if club-facilitated local interlocks are established to help firms cope with resource dependence, we should have observed a positive association between resource dependence and local interlocking. Further, this predicted association should have been more pronounced when the local upper-class club measure was removed from our models, indicating that local upper-class clubs mediate the effect of dependence on interlocking. We might even have observed a positive interaction effect of local upper-class clubs and resource dependence on interlocking, indicating that local clubs facilitate the creation of interlocks motivated by dependence. However, the results reported in tables 5 and 6, as well as supplemental analyses (available on request), reveal none of this.

This suggests that club-facilitated local interlocks are social class phenomena, reflecting the status concerns of corporate elites rather than the

organizational interests of the corporations they command. Exclusive social clubs may be the big city upper-class equivalent of the small town working-class bar—providing a venue for elites to become acquainted and intimate with one another. But, they may also be the medium through which a bounded local upper class is formed (Domhoff 1970). If so, local upper-class formation is likely driven by individual elite interests in advancement, privilege, and the public display of membership in a closed status group. This accords well with our unexpected finding that high debt-to-equity ratios reduce local financial interlocking. The directors of failing corporations may obtain few invitations to sit on the boards of other local firms because only directors of successful firms are allowed to enjoy and display upper-class status.

Local interlocking, especially local financial interlocking, is also uniquely a function of two dimensions of the geography of corporate production. Local growth dependence increases local interlocking, while local economic dominance reduces it. A corporation's interests and power vis-à-vis local political processes that shape local growth systematically affect its participation in the local interlock network. Thus the local interlock network is a corporate infrastructural component of the growth machine. While the growth machine thesis has enjoyed widespread acceptance among community power researchers, this is the first systematic exposition of the implications this thesis holds for corporate interlocking.

Finally, local industrial interlocking is a function of local intercorporate ownership. However, our results indicate that local intercorporate ownership relations operate primarily at the upper end of the interlock range—distinguishing firms that maintain no local interlocks from those that maintain many. Examination of the biggest linkers in our study indicates that the effects of intercorporate ownership disproportionately reflect the tendency of high levels of intercorporate ownership to be associated with substantial interlocking. This is typified by four Kaiser family firms headquartered in Oakland, California—all linked by high levels of intercorporate ownership (between 42% and 79%) and all tied by large numbers of interlocks (between 24 and 26). However, many of the reported intercorporate ownership effects persist even when majority-owned corporations, such as the four Kaiser family firms, are eliminated from our sample—suggesting that even more modest levels of intercorporate ownership are associated with elevated levels of interlocking (results available on request).

As noted above, local interlocking is not related to resource dependence relations emphasized by past researchers—whether industrial or financial. This null result is important in its own right. First, it suggests that regionalization of the corporate interlock network is not the result of regionalized resource dependence relations. Perhaps this should not be surprising, inso-

far as most firms situated a substantial proportion of their production facilities outside their headquarters state by the 1960s (Friedland, Palmer, and Stenbeck 1990). By this time competition and exchange among production facilities had become national and even global (Barnet and Muller 1974). Second, it suggests that bank centrality in regional networks is not a consequence of localized financial dependence and thus that bank centrality in the national network is not the aggregate outcome of a multiplicity of regionalized financial dependence relationships. If debt dependence is a source of bank centrality in regional clusters, it is a generalized dependence experienced by all firms, which operates regardless of a firm's particular reliance on loan capital.

Understanding nonlocal interlocking.—Nonlocal interlocking, as distinct from local interlocking, is a product of being headquartered in a city with few financial headquarters and no upper-class clubs.¹² This might simply reflect the fact that corporations headquartered in cities with upper-class clubs and many financial headquarters need look no further than locally situated elites to find qualified candidates to fill board vacancies. Alternatively, being headquartered near financial headquarters and upper-class clubs might be a functional substitute for nonlocal interlocking. Being headquartered in a financial center might provide corporations with other means to obtain information about, access to, and control over national capital flows and might provide nonlocal financial institutions with other means to obtain information about and influence over the investment decisions of local firms. Corporations might even situate their headquarters proximate to financial headquarters and upper-class clubs in order to gain information about and influence over firms headquartered in distant locales upon which they depend. Economic geographers maintain that major financial centers provide corporations with strategic information about the national and global economies (Pred 1977; Cohen 1981; Noyelle and Stanback 1984). Further, Domhoff (1970) and Ratcliff (1980a, 1980b) maintain that local upper-class clubs integrate large firms into the national economy. However, it would be surprising if proximity to financial headquarters and upper-class clubs operates as a substitute for, as opposed to a complement or facilitator of, nonlocal interlocking. Functional alternatives to interlocking available to corporations headquartered near financial headquarters, which in turn may repre-

¹² Our upper-class club results run counter to those of Johnson and Mintz (1989), which imply that local clubs provide vehicles for local directors to associate with nonlocal directors and consequently to serve on the same corporate boards. Perhaps this discrepancy stems from the fact that our study and their study have very different purposes and methodologies. While we investigate links between corporations and include only directional interlocks, they examine links between directors and include only multiple connections.

sent lending consortia composed of nonlocal financials, include specialized business consulting enterprises, common neighborhood residences, shared civic group associations, and informal social gatherings.¹³

Nonlocal as opposed to local interlocking is also a function of geographic, competitive, and financial resource dependence. There are two possible explanations of this result. Corporations might tend to situate their headquarters far from those of the firms upon which they rely for resources. If so, as their dependence on other firms increased, corporations would be compelled to increase nonlocal rather than local interlocking. This is unlikely. While heavily indebted firms probably must look beyond locally headquartered financial institutions to secure sufficient supplies of loan capital, we know that the financial institutions designated by corporations as their primary bank and the lead bank of any bank consortia to which they are indebted tend to be headquartered locally (Friedland and Palmer 1994). We also know that the tendency of corporations producing in the same industry to colocate (locate their headquarters in the same city) is quite high in the United States (Gardinali et al. 1996). More likely, being headquartered near other industrial and financial headquarters, and the low-cost social interaction it affords, is a functional substitute for local interlocking. In an earlier study, we found that interlock partners involved in some form of intercorporate coordination, such as a joint venture or long-term contract, were less likely to renew "accidentally disrupted" interlock connections if they were headquartered in the same city (Palmer et al. 1986; see also, Burt 1986a). Corporations might even colocate their headquarters strategically in response to market uncertainty and constraint. There is evidence that interindustry variation in headquarters colocation levels is associated with interindustry variation in market concentration (Gardinali et al. 1996).

The significance of exclusive upper-class social clubs.—The pervasive and in some respects surprising association between the presence of exclusive upper-class clubs in a city and the extent to which corporations headquartered there maintain local and nonlocal interlocks raises other questions. In particular, is the presence of exclusive upper-class clubs in a city a consequence of or correlated with other social processes? And if so, are

¹³ Indeed, the number of financial institutions and industrial firms headquartered in a city are highly correlated ($r = .99$). And prior research indicates that the number of specialized business service enterprises in a city is a function of the number of corporate headquarters situated there and the strength of this association is negatively related to the level of nonlocal interlocking typical of the locally headquartered firms (Palmer et al. 1990). Further, the largest corporations, presumably in need of the greatest amount of strategic information, tend to locate in the largest corporate headquarters complexes, presumably the richest sources of strategic information (Gardinali, Friedland, and Martinotti 1996).

these other social processes the more proximate determinants of interlocking?

Clearly, the formation of exclusive upper-class clubs in a city is contingent on the local residence and ascendance of upper-class individuals. And in 20th-century America, this in turn has been predicated primarily on the coincidental siting and development of capitalist enterprises—in particular, the administrative components of large industrial corporations and financial institutions. The formation of upper-class clubs is also contingent on the perception of local upper-class individuals that they belong to a distinct and closed social group. Thus, it is not surprising that corporate headquarters complexes with exclusive upper-class clubs have more industrial corporate and financial institution headquarters than do cities without such clubs. Further, upper-class club cities are more likely than nonclub cities to have a social register, which lists local upper-class individuals and their defining characteristics (e.g., prep school and college attendance, club memberships, and summer addresses) and thus presumes the existence of a local upper class as a self-conscious entity (table 7).

Do the preconditions of upper-class club formation, rather than the presence of upper-class clubs *per se*, shape interlocking? We cannot answer this question definitively here. Insofar as the establishment of upper-class clubs is likely to be integrally related—as both consequence and cause—to upper-class formation, it may be difficult to answer this question with future research as well. However, our results indicate that the association between the presence of upper-class clubs and corporate interlocking in a city is not simply a reflection of a deeper relationship between the presence of local headquarters and interlocking. Our club effects persist, and indeed even dominate, when the number of industrial and financial headquarters in a corporation's headquarters city are controlled. Indeed, they persist, as do our other results, even when a dummy variable indicating whether or not firms are headquartered in either of the two largest headquarters complexes—New York City and Chicago—is included in the analysis (results available on request).

Do other social processes, correlated with the establishment of upper-class clubs in a city, explain the observed club effects? One might surmise that upper-class clubs tend to be situated in particular regions—regions with unique cultural predispositions to interlocking. To examine this possibility, we examined the association between a city's region and the likelihood that it contained an upper-class club (tables 8 and 9). Green (1981a, 1981b) argues that firms headquartered in the South maintain more geographically dispersed interlock patterns and speculates that this partly accounts for the South's recent dramatic economic growth. Might our upper-class club effects actually signify Southern city effects? We doubt it. Cities with upper-class clubs are as likely to be located in the South

TABLE 7

CORPORATE HEADQUARTERS COMPLEXES WITH AND WITHOUT SOCIAL CLUBS

Corporate Headquarters Complex	No. of Industrial Corporate Headquarters	No. of Financial Institution Headquarters	Existence of Social Register
With a club:			
Seattle, Wash.	1	4	+
Providence, R.I.	2	0	
Richmond, Va.	2	2	
Buffalo, N.Y.	2	3	+
Houston, Tex.	3	3	+
Denver, Colo.	4	0	
Portland, Oreg.	4	2	
Dallas, Tex.	8	4	
Boston, Mass.	8	5	+
Minneapolis, Minn.	9	5	
Milwaukee, Wis.	10	4	
San Francisco, Calif.	10	7	+
Detroit, Mich.	12	4	+
St. Louis, Mo.	12	6	+
Los Angeles, Calif.	13	9	+
Philadelphia, Pa.	14	10	+
Pittsburgh, Pa.	22	2	+
Cleveland, Ohio	22	3	+
Chicago, Ill.	48	11	+
New York City	171	67	+
Without a club:			
Birmingham, Ala.	2	1	
Beverly Hills, Calif.	3	5	
Indianapolis, Ind.	4	0	
Fort Wayne, Ind.	4	1	
Rochester, N.Y.	5	0	
Akron, Ohio	5	0	
Toledo, Ohio	5	0	
Oakland, Calif.	5	0	
Wilmington, Del.	5	1	
Cincinnati, Ohio	5	5	+

NOTE.—Cities are considered corporate headquarters complexes if at least two MACNET industrial corporate or financial institutions were headquartered there in 1964. Eight cities possessing an exclusive upper-class club had fewer than two MACNET firm headquarters—New Orleans, Charleston, Palm Beach, Hartford, Baltimore, New Haven, Atlanta, and San Diego. Two cities possessing a social register had fewer than two MACNET firms—Baltimore and New Orleans.

TABLE 8

REGION OF CITIES WITH AND WITHOUT EXCLUSIVE UPPER-CLASS CLUBS

	No of Cities	% East	% Midwest	% West	% South
Club City	121	28.6	23.8	28.6	19.0
Non-Club City	136	37.3	38.7	9.3	14.7

as they are to be headquartered elsewhere. More generally, one might guess that upper-class clubs tend to be situated in older Eastern cities, which are bastions of inbred business practices. If so, our upper-class club effects might really proxy Eastern city effects. But this too is doubtful, insofar as cities with upper-class clubs are actually less likely than cities without such clubs to be situated in the East. If anything, our club city effects would have to represent Western or Midwestern city effects. However, there is no theory or conventional wisdom that would predict such an effect.

Finally, one might surmise that upper-class clubs tend to be located in cities where particular industries predominate—industries in which less aggressive and spatially expansive environmental monitoring and management strategies are prevalent. Such strategies might reflect the older age (i.e., position in the product life cycle) or other characteristics (i.e., inputs and outputs) of the industries in question. To evaluate this possibility, we grouped each of our firms into one of 10 primary industry categories suggested by Schmenner (1982). The categories capture the extent to which firms producing in them face different locational imperatives, partly as a consequence of their age and position in the product life cycle, and thus possibly possess different needs to maintain local and nonlocal interfirm ties. On the one hand, firms specializing in agriculture, forest, and mining industries might have more geographically dispersed organizational interests because their inputs are more dispersed in space. Those producing in market sensitive and labor-rate sensitive industries might also tend to have more dispersed organizational interests because they must maximize access to dispersed consumer and labor markets. On the other hand, firms specializing in industrial machinery/transportation equipment, high technology, and specialty chemicals might have relatively geographically circumscribed interests because they are dependent on agglomerative economies. Firms producing in two industries—heavy metals and heavy chemicals—experience contradictory locational pressures (Friedland, Palmer, and Stenbeck 1990). However, cities with upper-class clubs do not differ noticeably from cities without clubs, in

TABLE 9
PRIMARY INDUSTRY OF FIRMS IN CITIES WITH AND WITHOUT EXCLUSIVE UPPER-CLASS CLUBS

	No of Firms	% Agriculture Total	% Market Sensitive	% Forest Total	% Labor-Force Sensitive	% Heavy Chemicals	% Speciality Chemicals/ Metals	% Heavy Metals	% Industrial Machinery	% High Tech	% Mining
Club City	295	10.2	14.2	4.4	8.8	14.9	7.8	11.9	20.3	3.7	2.7
Non-Club City	179	8.9	5.0	5.0	12.3	16.8	8.4	8.4	24.0	8.9	1.7

regards to the primary industries in which the corporate headquarters they contain produce (table 9). Thus, for the time being, we conclude that exclusive local upper-class social clubs have an independent effect on interlocking.

Implications, Limitations, and Future Research

These speculations have a number of theoretical and practical implications. First, the local capitalist class social organization made possible by the colocation of corporate headquarters and upper-class clubs may be an instrument by which corporate elites manage their organizational environments, gaining information, trusted contacts, and access to national networks of elites. However, this local capitalist class organization may not serve organizational interests *through the creation of interlocks*. Second, the structure of the interlock network may be as much a result of the relative headquarters locations of resource dependent corporations as it is a consequence of the magnitude of the resource dependence relationships between firms per se. In particular, bank centrality in the national network may not be the result of financial dependence in general. Rather it might stem from the geographic disjuncture between the headquarters of debt-dependent industrials and their financial intermediaries. Third, previous researchers may have underestimated the association between resource dependence and interlocking because they have lumped together nonlocal interlocking, motivated by the desire to manage resource dependencies, and local interlocking, superfluous to the conduct of such relations. Zajac's (1988) argument that interlocks are seldom created to manage resource dependence may only be true for local interlocking. Fourth, future researchers should consider corporate headquarters location as a potential strategy for the management of market uncertainty.

There is some evidence to support these speculations. If the desire to manage resource dependence motivates corporations to locate their headquarters in cities where many other industrial firms and financial institutions are headquartered and where upper-class clubs are situated, then measures of resource dependence should be associated with being headquartered in such cities. The bivariate correlations reported in table 2 show that large corporations maintaining geographically complex production structures, which situate a high percentage of their plants outside their headquarters city in many different states, tend to be headquartered in cities with numerous industrial and financial headquarters and at least one upper-class club. They also show that diversified corporations tend to be headquartered in cities with an upper-class club.

Nevertheless, there are at least three noticeable limitations of our study. First, we failed to measure the spatial structure of resource dependence

relations. Competition and resource exchange undoubtedly have a geographical as well as an industrial structure. Further, while the industrial structure of market relations is very stable (Burt 1988), their territorial structure is probably quite dynamic (Bluestone and Harrison 1982; Harrison 1994; Romo and Schwartz 1995). Ultimately, we need to develop measures that reflect the extent to which a corporation's production facilities and administrative offices are situated close to or distant from those of its actual and potential competitors, buyers, and suppliers. Then we need to theorize associations between these measures and interlocking. We have taken a tentative first step in this direction by speculating from our results that corporations headquartered proximate to the headquarters of firms upon which they are dependent have less need to interlock with them because the colocation of headquarters allows the elites of interdependent firms to interact with each other outside the board room. Much more remains to be done along these lines, though, if we are to advance our understanding of the relationships between resource dependence and corporate interlocking. In particular, because most large corporations are highly diversified, future attempts to measure the geography of a corporation's resource dependence relationships need to take into account the full range of industries—not just the primary industry—in which it produces. If we do not, even the most precise measurements of the geography of resource dependence relationships may be too crude to apprehend the relationship between dependence and interlocking.¹⁴

Second, our data pertain to the 1960s. Our earlier remarks not withstanding, technological innovations such as facsimile transmission, teleconferencing, and high-speed computer networking, which have further propelled the globalization of economic activity, may have minimized or altered the significance of the spatial processes uncovered here. Finally, we have only analyzed cross-sectional associations. Thus, our ability to draw causal inferences from our data is severely restricted.

We have shown that the spatial structure of corporate relations influences the extent and spatial character of interlocking. Does the spatial structure of interlocking influence corporate behavior? One study suggests that spatially dispersed interlock ties channel management consulting and advertising contracts out of a firm's headquarters city to enterprises in other locales (Palmer et al. 1990). We suspect that the spatial structure

¹⁴ We calculated a series of measures in which indicators of the resource dependence to which a corporation was subject (e.g., the mean-deviated concentration level in a firm's primary industry) were weighted by the extent to which the headquarters of other firms upon which it was resource dependent were located in its headquarters city (e.g., the number of firms producing in its primary industry that were also headquartered in its headquarters city). However, none of these measures were found to influence interlocking.

of corporate interlocking may also influence other aspects of corporate behavior, such as a firm's propensity to locate new production facilities in its headquarters locale, relocate its headquarters to new locales, and become involved in local and nonlocal acquisitions (Palmer and Friedland 1987). Further, if, as our results indicate, local and nonlocal interlocking have different antecedents, might not they have different consequences? To answer this question, we may need a better understanding of the processes that underpin the relationships interlocks—local and nonlocal alike—facilitate (Mizruchi 1996).

REFERENCES

- Allen, Michael Patrick. 1974. "The Structure of Interorganizational Elite Cooptation: Interlocking Corporate Directorates." *American Sociological Review* 39:393–406.
- . 1978. "Economic Interest Groups and the Corporate Elite Structure." *Social Science Quarterly* 58:597–615.
- Atwood, William, James Bearden, Peter Freitag, Carolyn Hendricks, Peter Mariolla, Beth Mintz, Mark Mizruchi, Donald Palmer, and Michael Schwartz. 1985. "Data Collection and Analysis for the Mathematical Analysis of Corporate Networks (MACNET)." Pp. 255–77 in *The Power Structure of American Business*, edited by Beth Mintz and Michael Schwartz. Chicago: University of Chicago Press.
- Baltzell, E. Digby. 1958. *Philadelphia Gentleman: The Making of a National Upper Class*. Glencoe, Ill.: Free Press.
- Barber, Brad, Donald Palmer, and James Wallace. 1995. "Determinants of Conglomerate and Predatory Acquisitions: Evidence from the 1960s." *Journal of Corporate Finance* 1 (1): 283–318.
- Barnet, Richard J., and Ronald E. Muller. 1974. *Global Reach: The Power of Multinational Corporations*. New York: Simon & Schuster.
- Barnett, William P., and Glenn Carroll. 1987. "Competition and Mutualism among Early Telephone Companies." *Administrative Science Quarterly* 32 (3): 400–421.
- Bearden, Jim, and Beth Mintz. 1985. "Regionality and Integration in the United States Interlock Network." Pp. 234–49 in *Networks of Corporate Power*, edited by F. Stokman, R. Ziegler, and J. Scott. Cambridge: Polity Press.
- . 1987. "The Structure of Class Cohesion: The Corporate Network and Its Dual." Pp. 187–207 in *Intercompany Relations: The Structural Analysis of Business*, edited by M. Mizruchi and M. Schwartz. Cambridge: Cambridge University Press.
- Bell, Daniel. 1973. *The Coming of Post-Industrial Society*. New York: Basic.
- Bluestone, B., and B. Harrison. 1982. *The Deindustrialization of America*. New York: Basic.
- Bonacich, Phillip, and G. William Domhoff. 1981. "Latent Classes and Group Membership." *Social Networks* 3:175–96.
- Bourdieu, Pierre. 1990. *The Logic of Practice*. Stanford, Calif.: Stanford University Press.
- Burch, Philip. 1972. *The Managerial Revolution Reassessed*. Lexington, Mass.: Heath.
- Burt, Ronald. 1980. "Coaptive Corporate Actor Networks: A Reconsideration of Interlocking Directorates Involving American Manufacturing." *Administrative Science Quarterly* 25:557–82.
- . 1983. *Corporate Profits and Cooptation*. New York: Academic Press.
- . 1986a. "Broken Ties and Corporate Markets." Manuscript. Columbia University, Department of Sociology.

- . 1986b. "Data on Aggregate American Markets during the 1960s and 1970s." Center for the Social Sciences Technical Report #TR4. Columbia University, Research Program in Structural Analysis.
- . 1988. "The Stability of American Markets." *American Journal of Sociology* 94:356–95.
- Burt, Ronald, Kenneth Christman, and Harold C. Kilburn. 1980. "Testing a Structural Theory of Corporate Cooptation." *American Sociological Review* 45:821–41.
- Cameron, Colin, and Pravin Trivedi. 1986. "Econometric Models Base on Count Data: Comparisons and Applications of Some Estimators and Tests." *Journal of Applied Econometrics* 1:29–53.
- . 1990. "Regression Based Tests for Overdispersion in the Poisson Model." *Journal of Econometrics* 46:347–64.
- Cohen, Robert. 1981. "The Corporation and the Metropolis." Working paper. Conservation of Human Resources, Columbia University.
- Crenson, Mathew. 1971. *The Un-Politics of Air Pollution: A Study of Non-Decisionmaking in the Cities*. Baltimore: Johns Hopkins University Press.
- Dill, William. 1958. "Environment as an Influence on Managerial Autonomy." *Administrative Science Quarterly* 2 (4): 409–43.
- Domhoff, G. William. 1970. *The Higher Circles: The Governing Class in America*. New York: Random House.
- . 1974. *The Bohemian Grove and Other Retreats*. New York: Harper & Row.
- Dooley, Peter. 1969. "The Interlocking Directorate." *American Economic Review* 59: 314–23.
- Feld, Scott. 1981. "The Focused Organization of Social Ties." *American Journal of Sociology* 86:1015–35.
- Festinger, Leon. 1953. "Group Attraction and Membership." Pp. 92–101 in *Group Dynamics: Research and Theory*, edited by D. Cartwright and A. Zander. Evanston, Ill.: Row, Peterson.
- Festinger, Leon, Stanley Schachter, and Kurt W. Back. 1950. *Social Pressure in Informal Groups*. New York: Harper.
- Fitch, Robert, and Mary Oppenheimer. 1970. "Who Rules the Corporation?" *Socialist Revolution*, pt. 1, 1 (4): 73–108; pt. 2, 1 (5): 61–114; pt. 3, 1 (6): 33–94.
- Fortune*. 1963. *The Fortune Plant and Product Directory: The 1,000 Largest U.S. Industrial Corporations, 1963–1964*. New York: *Fortune*.
- . 1965. *The Fortune Directory: The 500 Largest U.S. Industrial Corporations*. New York: *Fortune*.
- Foucault, Michel. 1977. *Discipline and Punish*. New York: Pantheon.
- Friedland, Roger. 1983. *Power and Crisis in the City*. New York: Shocken.
- Friedland, Roger, and Deirdre Boden. 1994. Introduction to *NowHere: Space, Time, and Modernity*, edited by Roger Friedland and Deirdre Boden. Berkeley and Los Angeles: University of California Press.
- Friedland, Roger, and Donald Palmer. 1984. "Park Place and Main Street: Business and the Urban Power Structure." *Annual Review of Sociology* 10:393–416.
- . 1994. "Space, Corporation, and Class: Towards a Grounded Theory." Pp. 287–334 in *NowHere: Space, Time, and Modernity*, edited by Roger Friedland and Deirdre Boden. Berkeley and Los Angeles: University of California Press.
- Friedland, Roger, Donald Palmer, and Magnus Stenbeck. 1990. "The Geography of Corporate Production: Urban, Industrial, and Organizational Systems." *Sociological Forum* 5 (3): 335–59.
- Galaskiewicz, Joseph, Stanley Wasserman, Barbara Rauschenbach, Wolfgang Bielefeld, and Patti Mullaney. 1985. "The Influence of Corporate Power, Social Status, and Market Position on Corporate Interlocks in a Regional Network." *Social Forces* 64:403–31.
- Galbraith, John K. 1971. *The New Industrial State*. New York: Mentor.

- Gardinali, Paolo, Roger Friedland, and Guido Martinotti. 1996. "Corporate Headquarters Location: A Comparative Geography of Italy and the United States." Paper presented at the meeting of the American Sociological Association, New York City.
- Giddens, Anthony 1990. *Consequences of Modernity*. Stanford, Calif.: Stanford University Press.
- Glasberg, Davita Silfen 1989 *The Power of Collective Purse Strings: The Effects of Bank Hegemony on Corporations and the State*. Berkeley and Los Angeles: University of California Press.
- Green, Milford B 1981a. "The Corporate Interlocking Directorate as an Urban Spatial Information Network." *Urban Geography* 2 (2): 148-60.
- . 1981b. "Regional Preferences for Interlocking Directorates among the Largest American Corporations." *Environment and Planning* 3:829-39
- . 1983. "The Interurban Corporate Interlocking Directorate Network of Canada and the United States: A Spatial Perspective." *Urban Geography* 4 (4): 338-54.
- Greene, William. 1993 *Econometric Analysis*. New York: Macmillan.
- Harrison, Bennett. 1994. *Lean and Mean: The Changing Landscape and Corporate Power in the Age of Flexibility*. New York: Basic.
- Harvey, David. 1982. *The Limits of Capital*. Oxford: Basil Blackwell.
- . 1985. *The Urbanization of Capital*. Oxford: Basil Blackwell.
- . 1989. *The Condition of Post Modernity*. Oxford: Basil Blackwell.
- Haveman, Heather, and Lynn Nonnemaker. 1997 "Competition in Multiple Geographic Markets: The Impact on Market Entry and Growth." Manuscript Cornell University, Johnson School of Management.
- Haveman, Heather, and Elaine Romanelli. 1997. "Geography and Organizational Founding." Manuscript. Cornell University, Johnson School of Management
- Johnson, Eugene, and Beth Mintz. 1989 "Organization versus Class Components of Director Networks." In *Networks of Power: Organisational Actors at the National, Corporate, and Community Levels*, edited by Robert Perrucci and Harry R. Potter. New York: Aldine de Gruyter.
- Kadushin, Charles. 1995. "Friendship among the French Financial Elite." *American Sociological Review* 60 (2): 202-21.
- Kavaler, Lucy. 1960. *The Private World of High Society*. New York: David McKay.
- Koenig, Thomas, and Robert Gogel. 1981. "Interlocking Corporate Directorships as a Social Network." *American Journal of Economics and Sociology* 40:37-50.
- Koenig, Thomas, and John Sonquist 1977. "Interlocking Directorates in the Top U.S. Firms." *Insurgent Sociologist* 5:196-229.
- Kono, Clifford Tadao. 1990. "Weaving Spiders in Space: The Regionalization of Corporate Interlocks in the United States." Master's thesis. University of California, Santa Barbara, Department of Sociology.
- Korn/Ferry International. 1989. *Sixteenth Annual Board of Directors Survey*. New York: Korn/Ferry International.
- Kotz, David. 1978. *Bank Control of Large Corporations in the United States*. Berkeley and Los Angeles: University of California Press.
- Levine, Joel. 1972. "The Sphere of Influence." *American Sociological Review* 37:14-27.
- Lincoln, James 1977. "Organizational Dominance and Community Structure." Pp. 19-50 in *Power, Paradigms, and Community Research*, edited by Roland Liebert and Allen Imershein. Beverly Hills, Calif.: Sage.
- Lincoln, James, Michael Gerlach, and Peggy Tagahashi. 1992. "Keiretsu Networks in Japan: A Dyad Analysis of Intercompany Ties." *American Sociological Review* 57 (5). 561-85.
- Logan, John, and Harvey Molotch. 1987. *Urban Fortunes*. Berkeley and Los Angeles: University of California Press.

- Mace, Myles. 1986. *Directors: Myth and Reality*. Boston: Harvard Business School Press.
- Marquis, Inc. 1965. *Who's Who in Commerce and Industry*. Chicago: Marquis.
- Mills, C. Wright. 1956. *The Power Elite*. New York: Oxford University Press.
- Mintz, Beth, and Michael Schwartz. 1981a. "Interlocking Directorates and Interest Group Formation." *American Sociological Review* 46:851-68.
- . 1981b. "The Structure of Intercorporate Unity in American Business." *Social Problems* 29:87-103.
- . 1985. *The Power Structure of American Business*. Chicago: University of Chicago Press.
- Mizruchi, Mark. 1982. *The American Corporate Network, 1904-1974*. Beverly Hills, Calif.: Sage.
- . 1989. "Similarity of Political Behavior among Large American Corporations." *American Journal of Sociology* 95:401-24.
- . 1992. *The Structure of Corporate Political Action*. Cambridge, Mass.: Harvard University Press.
- . 1996. "What Do Interlocks Do? An Analysis, Critique, and Assessment of Research on Interlocking Directorates." *Annual Review of Sociology* 22:271-98.
- Mizruchi, Mark, and David Bunting. 1981. "Influence in Corporate Networks: An Examination of Four Measures." *Administrative Science Quarterly* 26:475-89.
- Mizruchi, Mark S., and Linda Brewster Stearns. 1988. "A Longitudinal Study of the Formation of Interlocking Directorates." *Administrative Science Quarterly* 33:194-210.
- Molotch, Harvey. 1976. "The City as a Growth Machine: Toward a Political Economy of Place." *American Journal of Sociology* 82 (2): 309-31.
- . 1979. "Capital and Neighborhood in the United States: Some Conceptual Links." *Urban Affairs Quarterly* 14 (3): 289-312.
- Moody's Investor Service. 1965. *Industrial Manual*. New York: Moody's Investor Service.
- Mruk, Edwin S., and James A. Giardina. 1983. *Organization and Compensation of Boards of Directors*. New York: Arthur Young.
- Noyelle, Thierry, and Thomas Stanback. 1984. *The Economic Transformation of American Cities*. Totowa, N.J.: Rowman & Allanheld.
- Ornstein, Michael D. 1984. "Interlocking Directorates in Canada: Intercorporate or Class Alliance?" *Administrative Science Quarterly* 29:210-31.
- Palmer, Donald. 1983. "Broken Ties: Interlocking Directorates and Intercorporate Coordination." *Administrative Science Quarterly* 28:40-55.
- Palmer, Donald, Brad M. Barber, Xueguang Zhou, and Yasemin Soysal. 1995. "The Friendly and Predatory Acquisition of Large U.S. Corporations in the 1960s: The Other Contested Terrain." *American Sociological Review* 60 (4): 469-99.
- Palmer, Donald, and Roger Friedland. 1987. "Corporation, Class, and City Growth." Pp. 145-84 in *Structural Analysis of Business*, edited by Mark Mizruchi and Michael Schwartz. Cambridge: Cambridge University Press.
- Palmer, Donald, Roger Friedland, Amy Roussel, and P. Devereaux Jennings. 1990. "Corporations and the Urban Advanced Business Service Sector." *Social Forces* 69 (1): 115-37.
- Palmer, Donald, Roger Friedland, and Jitendra Singh. 1986. "The Ties that Bind: Organizational and Class Bases of Stability in a Corporate Interlock Network." *American Sociological Review* 51:781-96.
- Palmer, Donald, P. Devereaux Jennings, and Xueguang Zhou. 1993. "Late Adoption of the Multidivisional Form by Large U.S. Corporations: Institutional, Political, and Economic Accounts." *Administrative Science Quarterly* 38:100-131.
- Pennings, Johannis. 1980. *Interlocking Directorates*. San Francisco: Jossey-Bass.
- Perlo, Victor. 1957. *The Empire of High Finance*. New York: International.

- Perrucci, R., and M. Pilisuk. 1970. "Leaders and Ruling Elites: The Interorganizational Basis of Community Power." *American Sociological Review* 36: 1040-57.
- Pfeffer, Jeffrey. 1972. "Size and Composition of Corporate Boards of Directors: The Organization and Its Environment." *Administrative Science Quarterly* 17:218-28.
- . 1992. *Managing with Power*. Boston: Harvard Business School Press.
- Pfeffer, Jeffrey, and Gerald R. Salancik. 1978. *The External Control of Organizations*. New York: Harper & Row.
- Phillips, Peter. 1994. "Relative Advantage: Sociology of the San Francisco Bohemian Club." Ph.D. thesis. University of California, Davis, Department of Sociology.
- Pred, Alan. 1977. *City Systems in Advanced Economies*. New York: Wiley.
- Ratcliff, Richard. 1980a. "Banks and Corporate Lending: An Analysis of the Impact of the Internal Structure of the Capitalist Class on the Lending Behavior of Banks." *American Sociological Review* 45:553-70.
- . 1980b. "Declining Cities and Capitalist Class Structure." Pp. 115-18 in *Power Structure Research*, edited by G. William Domhoff. Beverly Hills, Calif.: Sage.
- Romo, Frank, and Michael Schwartz. 1995. "The Structural Embeddedness of Business Decisions: The Migration of Manufacturing Plants in New York State." *American Sociological Review* 60 (6): 874-907.
- Schmenner, Roger. 1982. *Making Business Location Decisions*. Englewood Cliffs, N.J.: Prentice Hall.
- Soja, Edward. 1989. *Postmodern Geographies*. London: Verso.
- Soref, Michael. 1976. "Social Class and a Division of Labor within the Corporate Elite: A Note on Class, Interlocking, and Executive Committee Membership of Directors of U.S. Industrial Firms." *Sociological Quarterly* 17:360-68.
- . 1980. "The Finance Capitalists." In *Classes, Class Conflict, and the State*, edited by Maurice Zeitlin. Cambridge, Mass.: Winthrop.
- Soref, Michael, and Maurice Zeitlin. 1987. "Finance Capital and the Internal Structure of the Capitalist Class in the United States." Pp. 56-84 in *Intercompany Relations: The Structural Analysis of Business*, edited by M. Mixruchi and M. Schwartz. Cambridge: Cambridge University Press.
- Standard and Poor's. 1965. *Register of Corporations, Directors, and Executives*. New York: Standard & Poor's.
- Stanworth, Philip, and Anthony Giddens. 1975. "The Modern Corporate Economy: Interlocking Directorates in Britain, 1906-1970." *Sociological Review* 23:5-28.
- Stearns, Linda Brewster, and Mark S. Mixruchi. 1987. "Broken Tie Reconstitution and the Functions of Interorganizational Interlocks: A Reexamination." *Administrative Science Quarterly* 31:522-38.
- . 1993. "Board Composition and Corporate Financing: The Impact of Financial Institution Representation on Borrowing." *Academy of Management Journal* 36: 603-18.
- Sweezy, Paul. 1953. "Interest Groups in the American Economy." Pp. 62-82 in *The Past as Present*, edited by Paul Sweezy. New York: Monthly Review Press.
- U.S. House. 1968. Committee on Banking and Currency, Pattman Subcommittee on Domestic Finance. *Commercial Banks and Their Trust Activities: Emerging Influence on the American Economy*. New York: Arno Press, New York Times.
- Useem, Michael. 1979. "The Social Organization of the American Business Elite and Participation of Corporate Directors in the Governance of American Institutions." *American Sociological Review* 44:553-72.
- . 1984. *The Inner Circle. Large Corporations and the Rise of Business Politics in the U.S. and U.K.* New York: Oxford University Press.
- . 1987. "The Inner Circle and the Political Voice of Business." Pp. 143-62 in *The Structure of Power in American Business*, edited by Michael Schwartz. New York: Holmes & Meier.

- Useem, Michael, and Jerome Karabel. 1986. "Pathways to Corporate Management." *American Sociological Review* 51:184-200.
- Whitten, Guy D., and Harvey D. Palmer. 1996. "Heightening Comparativists' Concern for Model Choice: Voting Behavior in Great Britain and the Netherlands." *American Journal of Political Science* 40 (1): 231-60.
- Zajac, Edward J. 1988. "Interlocking Directorates as an Interorganizational Strategy." *Academy of Management Journal* 31:428-38.
- Zeitlin, Maurice. 1974. "Corporate Ownership and Control: The Large Corporation and the Capitalist Class." *American Journal of Sociology* 81:894-903.
- Zeitlin, Maurice, Richard Earl Ratchiff, and Lynda Ann Ewen. 1974. "The Inner Group: Interlocking Directorates and the Internal Differentiation of the Capitalist Class in Chile." Paper presented at the meeting of the American Sociological Association, Montreal.

Caveat Emptor: The Construction of Nonprofit Consumer Watchdog Organizations¹

Hayagreeva Rao
Emory University

This article investigates how new organizational forms are constituted as cultural objects. Since new organizational forms jeopardize existing interests, institutional entrepreneurs recombine prevalent cultural materials to frame the form as necessary, valid, and appropriate. When rival entrepreneurs promote incompatible frames, the frame that enjoys greater political support from the state, professions, and other organizations becomes ascendant. Proponents of losing frames can exit, migrate, or convert to the ascendant frame. A case study of the creation of nonprofit consumer watchdog organizations demonstrates how the boundaries of an organizational form and its cultural contents are shaped by politics.

Where new organizational forms come from is one of the central questions of organizational theory. New organizational forms are new embodiments of goals, authority, technology, and client markets. Novel social structures matter because they underpin organizational diversity. The ability of societies to respond to social problems hinges on the diversity of organizational forms, and in the long run, in a fluid environment, diversity can be maintained or increased by the rise of new forms (Hannan and Freeman 1989, p. 3). Moreover, new forms are consequential motors of evolution—indeed, organizational change consists of the replacement of existing organizational forms by new organizational forms (Schumpeter 1950; Astley 1985). Furthermore, since new forms are structural incarnations of beliefs, values, and norms, they emerge in tandem with new institutions and foster cultural change in societies (Durkheim 1950, p. 16; Stinchcombe 1965; Scott 1995).

It is only recently that organizational theorists have begun to analyze

¹ I thank Cynthia Hewitt, Young Chang, Robert Kadel, and Farah Mihoubi for their excellent research support. I am grateful to Howard Aldrich, Joel Baum, Sadhna Diwan, John W. Meyer, Royston Greenwood, and Sidney Winter for valuable suggestions. Direct correspondence to Hayagreeva Rao, Roberto C. Goizueta Business School, Emory University, Atlanta, Georgia 30322. E-mail: H_rao@bus.emory.edu

the origins of new organizational forms from the standpoint of the random variation, constrained variation, and the cultural-frame institutional perspectives. The *random variation perspective* is premised on biological evolutionary models and holds that realized variations in organizational forms are random. Proponents of this view suggest that new forms arise when search routines lead to modifications of operating routines (Nelson and Winter 1982), or when a small group of competence-sharing organizations is isolated and finds a favorable resource environment (McKelvey 1982). Despite its appeal, the random variation perspective is of limited use since it is difficult to empirically identify competencies and routines. Another major drawback is that it is silent on the specific processes that generate variations and on the content of variations.

The *constrained variation perspective* asserts that conditions in the environment predictably foster or diminish variations in organizational forms. Its different versions emphasize creative destruction through technological innovation (Schumpeter 1950; Tushman and Anderson 1986), environmental imprinting, wherein social conditions at the time of founding limit organizational inventions (Stinchcombe 1965; Kimberly 1975), and existing organizations as producers of new organizational forms (Brittain and Freeman 1982; Lumsden and Singh 1990). A key premise common to all versions of the constrained variation perspective is that the existence of ecospace unoccupied by other forms is an important precondition for the birth of new organizational forms. Nevertheless, proponents of the constrained variation perspective differ with respect to the antecedents of resource spaces. If models of creative destruction hold that the demise of existing organizations frees resources for new organizations, then models of environmental imprinting stress the importance of political upheavals (Carroll, Delacroix, and Goodstein 1990), entrepreneurs' access to wealth and power, labor markets, and the protective role of the state (see Aldrich 1979). By contrast, those who portray existing organizations as producers of new organizations hold that interrelations among existing organizations influence the branching of new resource spaces (Carroll 1984) and the ability of existing organizations to exploit new resource spaces (Romanelli 1989). Although the constrained variation perspective usefully emphasizes the primacy of resource spaces, its drawback is that resources do not preexist as pools of free-floating assets but have to be mobilized through opportunistic and collective efforts (Van de Ven and Garud 1989). Another defect is that the constrained variation perspective elides how formal structures become imbued with norms, values, and beliefs during the process of resource mobilization by entrepreneurs.

The *cultural-frame institutional perspective* complements the constrained variation perspective and proposes that new organizational forms arise when actors with sufficient resources see in them an opportunity to

realize interests that they value highly.² A core premise is that the creation of all new organizational forms requires an institutionalization project wherein the theory and values underpinning the form are legitimated by institutional entrepreneurs (DiMaggio 1988, p. 18). In this perspective, institutional projects can arise from organized politics or social movements. When they arise from organized politics, they still also resemble social movements in that the resources and interests of actors are not fixed and the rules governing interaction are contested (Padgett and Ansell 1993; Fligstein 1996a, p. 5). The cultural-frame institutional perspective on the rise of new organizational forms is, as yet, emergent and needs to "pay more attention to preexisting institutional conditions, what the alternative institutional projects are in a given situation, and the political process by which projects win out" (Fligstein 1996a, p. 27) and to "direct attention to the role of professionalization in the institutionalization of new forms, and more generally, to the establishment of fieldwide environments around the forms" (DiMaggio 1991, p. 289).

These gaps in the cultural-frame institutional perspective supply the motivation for this article to analyze the institutional production of nonprofit consumer watchdog organizations (CWOs) in the United States. CWOs are social control specialists who institutionalize distrust of agents by inspecting their performance on behalf of their clients and who sustain social order (Luhmann 1979; Zucker 1986; Shapiro 1987). For-profit CWOs like Morningstar and Lipper Analytical Services inform customers by disseminating evaluations of mutual funds. By contrast, nonprofit CWOs such as Consumers Union and the Council for Responsible Genetics evaluate products and technologies, lobby legislators and governmental agencies, and educate consumers about their rights. Unlike for-profit CWOs, nonprofit CWOs do not distribute surpluses as profits to shareholders.

Nonprofit CWOs are a significant outcropping of the social organization of distrust.³ Since the formation of the first nonprofit CWO, Consumers

² Cultural-frame institutionalist arguments differ from rational-choice institutional arguments. The latter imply that new organizational forms emerge when it is in the private interests of individuals to establish them and hold that "actors must establish a set of obligations as well as a mechanism that enforces compliance to these obligations" (Hechter 1990, p. 14). Institutional building episodes are bargaining games wherein actors with fixed preferences enter into bargains over issues that foster cooperation (Shepsle 1990). By contrast, cultural-frame institutional arguments assert that preexisting institutions and organizations constrain actors to cooperate (March and Olsen 1989). Moreover, the interests of actors are not fixed and new conceptions of interests can emerge (Padgett and Ansell 1993).

³ The number of employees in a nonprofit CWO is not a credible proxy of its importance because CWOs are minimalist organizations with few employees. Their biggest resource is usually subscribers to magazines or donors. Some nonprofit CWOs publish

Research in 1927, the number of national-level nonprofit CWOs in the United States has surged to 200 in 1995 (Bykerk and Manney 1995). Some national nonprofit CWOs are "second-order associations" (composed of other organizations as members), focusing on lobbying legislators and government, whereas other national nonprofit CWOs are "first-order associations," consisting of individual members. In 1995, the largest second-order association, the Consumer Federation of America (started in 1967) consisted of 240 national and regional organizations whose total membership was more than 50 million members. The largest first-order association was Consumers Union with 4.9 million members (Bykerk and Manney 1995). By virtue of their monitoring, lobbying, and educational activities, national nonprofit CWOs have induced American business firms to create special-purpose departments concerned with consumer affairs. Indeed, a new professional group called the Society for Consumers Affairs Professionals in Business was set up in 1973, and by 1988, it had 750 large firms as members (Fornell 1988). Nonprofit CWOs coexist with watchdogs specializing in environmental, civil rights, child welfare, animal welfare, voter rights, and human rights sectors.⁴ In a study of the American nonprofit watchdog community, Meyer and Imig (1993) reported 196 national watchdogs that sought to influence legislation in the consumer, child welfare, animal rights, civil rights, and antipoverty sectors.⁵ Although watchdogs in these sectors perform specialized functions, there is some interpenetration. Thus, prominent environmental watchdogs also espouse consumer agendas (Bykerk and Manney 1995); for example, Greenpeace promotes the idea that consumers ought to be environmentally responsible by purchasing "green" goods and disciplines errant firms by organizing consumer boycotts.

Despite their importance, there is a dearth of research on the social origins of nonprofit CWOs. Sociologists have analyzed nongovernmental nonprofit watchdogs such as accrediting bodies (Wiley and Zald 1968) and professional societies (Briloff 1972; Freidson 1986) but have overlooked the cultural origins of nonprofit CWOs. The burgeoning literature

test bulletins, but the circulation of consumer test bulletins severely understates the extent of their use. For example, a Gallup poll in England revealed that the readership of the consumer magazine *Which?* was 3 million in 1967 even though the actual subscriptions totaled 434,000 (Thorelli and Thorelli 1974).

⁴ Luhmann (1979, p. 57) observes, "In practical terms, control over trust can only be exercised as someone's main occupation. Everybody else must rely on the specialist involved in such control."

⁵ Note that this study understates the number of national-level watchdogs by not including service providers. Additionally, Meyer and Imig (1993) also condensed national organizations that shared phone numbers, addresses, and boards of directors into one listing.

on consumerism glosses over how institutional dynamics shape the design of nonprofit CWOs. Brobeck (1990, p. xxxii) surveyed 943 articles on consumer organizations and concluded that there was insufficient empirical research on the cultural origins of nonprofit consumer organizations, their advocacy and interrelationships, and their linkages with other organizations. In their detailed review of consumption, Frenzen, Hirsch, and Zerillo (1994, p. 410) noted that the "processual linking of producers and consumers is a rich, though still unexplored, research arena for economic sociology, with available archival and contemporary data sources waiting to be mined." If variability in the social organization of distrust is to be understood, as Shapiro (1987) recommends, then the study of the cultural origins of nonprofit CWOs is essential. Below, I begin by elaborating cultural-frame institutional arguments on the rise of new forms, and thereafter, I present a detailed case study of the formation of Consumers Research and Consumers Union to demonstrate cultural-frame institutional arguments.

INSTITUTIONAL DYNAMICS AND THE RISE OF NEW FORMS

The starting point of cultural-frame institutionalism is that new forms do not arise automatically in resource spaces but have to be constructed from prevalent cultural materials. In turn, detailed attention to the process of the construction of new forms devolves into an analysis of institutional entrepreneurs (Granovetter 1994). Institutional entrepreneurs are ideological activists who combine hitherto unconnected beliefs and norms into an organizational solution to a problem (Becker 1963).

Entrepreneurs, Frames, and Legitimacy

Stinchcombe (1968, p. 194) asserts that the entrepreneurial creation of new forms "is preeminently a political phenomenon" because support has to be mobilized for the goals, authority structure, technology, and clients embodied in the new form. In some cases, resource spaces unoccupied by other forms may exist, but the existence of such unfilled resource spaces does not mean that "free floating" resources are easily available to potential entrepreneurs. Entrepreneurs have to assemble resources, to legitimate the new form, and to integrate it with the prevalent institutional order. In other cases, resource spaces for a new form may not exist, and entrepreneurs have to construct these spaces by defining opportunity, by identifying distinctive resources, and by prying them away from existing uses. Since entrepreneurs are trying to convince others to go along with their view, the formation of new industries and forms resembles social movements (Fligstein 1996*b*, pp. 663–64).

Social movement theorists propose that institutional entrepreneurs can mobilize legitimacy, finances, and personnel through the use of frames (McAdam, McCarthy, and Zald 1988). Frames define the grievances and interests of aggrieved constituencies, diagnose causes, assign blame, provide solutions, and enable collective attribution processes to operate (Snow and Benford 1992, p. 150).⁶ Thus, frames are theories that justify an organizational form—an incarnation of goals, authority, technology, and clients, as indispensable, valid, and appropriate. Frames are tools of unobtrusive influence and can shape people's "perceptions, cognitions, and preferences in such a way that they accept their role . . . either because they see or imagine no alternative to it, or because they view it as natural" (Lukes 1974, p. 24). In "much the same way that pictures are framed, questions and actions are framed, and the context in which they are viewed and discussed determines what gets done . . . setting the context is a critical strategy for exercising power and influence" (Pfeffer 1992, p. 202).

Institutional entrepreneurs create frames by selecting items from a pre-existing cultural menu (Meyer and Rowan 1977, p. 345). Swidler (1986, p. 277) suggests that a culture is not a "unified system that pushes action in a consistent direction. Rather it is more of a 'tool kit' or a repertoire from which actors select differing pieces for constructing lines of action." Douglas (1986) points out that bricolage is an important method by which entrepreneurs construct new cognitive models and formal structures. Thus, entrepreneurs can recombine elements from existing repertoires through imitation or consciously revise existing models on the basis of their training in other organizations.

Multiple Frames and Conflict

Politics become obtrusive when an unfilled resource space "calls forth and permits a range of definitions of the situation" (Zald and McCarthy 1980, p. 6) and when rival coalitions of issue entrepreneurs champion incompatible frames. Even as entrepreneurs may draw on a generalized Western cultural account (Meyer, Boli, and Thomas 1987) and justify their actions on the basis of the widely accepted myths of progress and justice, there is a wide scope for conflict over the practical implications of the Western cultural account in the construction of new organizational forms. As mod-

⁶ Fligstein (1996b, p. 658) says that frames are really "conceptions of control," that is, "understandings that structure perceptions of how a market works and . . . allow actors to interpret their world and act to control situations." "Conceptions of control are shared cognitive structures within and across organizations that have profound effects on organization design and competition" (Fligstein 1996b, p. 671).

erates plead for progress and radicals glorify justice and clash like tectonic plates, interorganizational conflict ensues. Which frame and its organizational embodiment should be chosen to define and organize an activity is a political question (DiMaggio 1994; Tarrow 1989). Friedland and Alford (1991, pp. 240–42) propose that the creation of new organizational forms unfolds at three levels of analysis, with “individuals competing and negotiating, organizations in conflict and coordination, and institutions in contradiction and interdependency. . . . We conceive of these levels of analysis as ‘nested,’ where organization and institution specify higher levels of constraint and opportunity for individual action.”

When multiple frames and forms vie with each other, why one form is chosen and why other roads are not pursued hinge on larger constellations of power and social structure (Brint and Karabel 1991, p. 346). Struggles to produce new meanings and new social structures are, therefore, the motors of cultural change in societies (Tarrow 1989), and these tussles unfold in an organizational field where the state and the professions play an important role (DiMaggio and Powell 1983).

The state constrains the creation of new forms as a collective actor and as an institutional structure (Campbell and Lindberg 1990). As a set of semiautonomous actors, agencies of the state grant charters, allocate finance or monopoly status, and impose taxes and regulatory controls. Baron, Dobbin, and Jennings (1986) showed that personnel offices arose as document providers because firms had to file “manning tables” enumerating manpower needs and jobs in response to employment stabilization policies instituted by the federal government during World War II. As an institutional structure, the structure of the state, whether it is fragmented or centralized, affects organizational structures. Moreover, the executive, legislative, and judicial arms of the state also serve as arenas within which conflicts within and among organizations are adjudicated. Finally, states have the capacity to define and enforce property rights, rules that determine the conditions of ownership and the control of the means of production (Campbell and Lindberg 1990, p. 635).

Professions shape the rise of new organizational forms by providing cognitive frameworks and by spawning formal structures to create and defend jurisdictional claims (Freidson 1986; Abbott 1990). Professionals can be conservatives in organizational roles and simultaneously use fieldwide organizations to reform the system that employs them. In the case of museums, administrative professionals dominated fieldwide structures and were at the forefront of debate about whether museums were to be structured like libraries, symphony orchestras, or department stores (DiMaggio 1991, p. 287). Professionals such as personnel experts and labor attorneys can also overstate legal threats (Sutton and Dobbin 1996) and can encourage organizations to institute defenses such as grievance proce-

dures. Alternatively, professionals who are sited outside organizations can impel organizations to conform to models dominant in the profession—thus, firms have been shown to dediversify when there is a mismatch between the industries in which the firm competes and the industry coverage of the financial analysts who follow it (Zuckerman 1996).

Apart from the state and the professions, the organizational field also consists of key suppliers, resource and product consumers, regulatory agencies, and other organizations that produce similar goods and services (DiMaggio and Powell 1983; DiMaggio 1994). Some of these actors may support a given frame, others may champion rival frames, and uncommitted organizations may withhold judgment and be neutral. It is in this multiorganizational field that institutional entrepreneurs interpret grievances, exchange evaluations, forge alliances, and joust with antagonists (Klandermans 1992). If alliances provide resources and opportunities for entrepreneurs, antagonists drain resources and restrict opportunities.

Boundaries of Forms

These arguments suggest that both entrepreneurs and frame conflict play a substantial role in bundling goals, authority structure, technologies, and clients into new organizational forms. Although organizational economists assert that technical considerations, especially transaction costs, determine “efficient boundaries” of organizational forms (Williamson 1985), institutional dynamics, as Hannan and Freeman (1989, pp. 43–59) note, play a double-edged role—they can blend differences among organizational forms or segregate activities into distinct organizational forms.

On the one hand, institutional processes can erode distinctions among organizational forms. Thus, bricolage can blur distinctions among organizational forms. Institutional entrepreneurs can recombine elements from existing repertoires by imitation into new theories or frames or by consciously redesigning existing theories of organization (Douglas 1986). Similarly, copying errors can lead to the accumulation of small unintended changes and can blur differences among forms as organizations diverge from common models (Powell 1991). Deinstitutionalization, wherein the rules defining social and technical boundaries are broken by the passage of legislation or by the crossover of personnel can also eliminate distinctions among forms (Oliver 1992).

Arguably, institutional processes are the “important segregating mechanism” (Hannan and Freeman 1989, p. 54). New forms become established when a powerful actor (e.g., the state) endorses the claims of social actors using a form, or when a form becomes taken for granted as the way to organize activities. Collective action by members of a group can also foster distinctions among forms (e.g., the formation of trade associations). More-

over, closure of social networks through inbreeding and replication (Powell 1991) can also encode distinctions among forms.⁷

Hannan and Freeman (1989, p. 60) observe that segregating and blending processes are not stages in an evolutionary process but are sub-processes that simultaneously edit organizational diversity. Rival bands of entrepreneurs can construct incompatible frames through bricolage and can promote structural proposals that are recombinations of preexisting organizational models. When rival blends vie for dominance, politics determine which frame is chosen and how boundaries are imposed on organizational activities. The success of collective action efforts and the endorsement of powerful actors shape the selection of a frame and the concomitant choice of organizational boundaries. Collective action and the endorsement of powerful actors become even more important when technical differences among rival frames and structural proposals are minimal. In cases where the criteria for a good technical solution are contested, political and institutional processes shape not only what organizations can do, but which organizations can exist (Powell 1991, pp. 186–87).

The coalition that garners the greatest political support will find that its frame will be privileged (Brint and Karabel 1991, p. 355; McAdam 1994). Whether a coalition wins or not hinges on its size, the existence of political opportunity, the attitude of state actors, its support from professionals, and its ability to build a political coalition around an identity (Fligstein 1996b, p. 664). Thus, the scope of the form (i.e., the goals, authority structure, technology, and clients subsumed by the form) are outcomes of contending attempts at control and competing quests to impose a preferred definition of the identity of the constituencies that benefit from the form (White 1992).

Nelson and Winter (1982, pp. 109–11) propose that intraorganizational routines become *operative* only when there is a comprehensive truce or when there is a cessation of conflict amongst members of an organization. Analogously, the boundaries of a new organizational form become established and the new form becomes integrated into an community of organizations only when there is a truce amongst the constituents of the organizational field about which frame is used to organize activities. Like truces among nations, truces among rival institutional entrepreneurs can also be

⁷ It is interesting that Hannan and Freeman (1989, p. 56) describe endorsement by powerful actors (such as the state) and the taken-for-granted character of forms as institutional processes but do not similarly characterize collective action (lobbying designed to create laws) and the closure of social networks (inbreeding). I suggest that all four are institutional processes.

unequal, with some winning a larger slice of the cake and as a result, having a privileged position for their frame.

March and Olsen (1989) suggest that conflicts may be resolved through the logic of aggregation and give and take, or the logic of integration wherein one of the parties can learn from the other and even convert to the other's point of view. Those who lose can exit the arena, concentrate on a different niche or even embrace the ascendant frame. When the proponents of a losing frame abandon deviant ideas and capitulate by adopting the ascendant frame, they can "normalize" themselves and become integrated into the social system (Hirsch 1986).

Truces increase the capacity for collective action by reducing comprehensiveness; some points of view are ignored or suppressed. The terms of a truce among rival institutional entrepreneurs can never be completely explicit, thus, the maintenance of truces depends upon the disincentives for actors for engaging in provocative actions and the defensive alertness of parties keen on preserving the status quo. As a result, just like intraorganizational routines, organizational forms are "confined to extremely narrow channels by the dikes of vested interest. Adaptations that appear 'obvious' and 'easy' to an external observer may be foreclosed because they involve a perceived threat to the . . . political equilibrium" (Nelson and Winter 1982, p. 111).

THE FORMATION OF NONPROFIT CWOS IN AMERICA

Historians of the consumer movement recognize three eras: the antiadulteration movement, culminating in the passage of the Pure Food and Drugs Act of 1906, the rise of nonprofit consumer watchdogs such as Consumers Research and the Consumers Union in the 1930s, and legal activism commencing with Ralph Nader's crusade for product safety in the 1960s (Hermann 1970).

This article focuses on the second wave of activism because it witnessed the rise of nonprofit CWOs as a new organizational form. The early history of a new organizational form, as Scott (1995, p. 147) notes, is a period of coevolution in which existing institutional arrangements are delegitimated and new norms and beliefs are being constructed. This article relies on a wide variety of archival and publicly available sources—documents of the Consumers Union, Consumers Research, newspapers, census statistics, and writings of issue entrepreneurs and historians. Historical research on the origins of an organizational form provides greater breadth than conventional ethnography and enables us to derive a historically informed understanding of organizations (DiMaggio 1991, 1994). In order to demonstrate how nonprofit CWOs emerged as a new organizational form, I begin

by discussing organizational precursors to nonprofit CWOs: standards and testing organizations and consumer leagues.

Precursor Organizations

Standards and testing organizations were an important organizational precursor to nonprofit CWOs and arose around 1900 as proponents of the idea of removing wasteful variety through standardization. By contrast, consumer leagues that arose in the early part of the 20th century sought to improve standards of living.

Standards and testing organizations.—Standards and testing organizations were mainly extensions of trade associations and professional societies seeking to promulgate common metrics to assist business organizations. In 1894, an association of insurance underwriters (Underwriters Laboratory) received a charter to certify wires and light fixtures as fire resistant in order to build insurable real estate. Other trade associations established standard nomenclatures and performance specifications in the wool blanket and laundry industries.

A professional society of engineers called the Society of Automotive Engineers (SAE) was established in 1906 for the explicit purpose of standardizing the supply of components to automobile manufacturers. Instead of having numerous suppliers making incompatible components, the SAE initially laid down 600 specific standards to ensure the interchangeability of components and thereby increased the size of the potential market for parts manufacturers and ensured reliability for auto producers (Chase and Schlink 1927). Similarly, materials testing experts promulgated standards for paint, and electrical engineers developed standards for electrical components for large business enterprises.

Standard-setting bodies that arose as extensions of trade associations and professional societies induced action by governmental authorities. The National Bureau of Standards (established in 1901) instituted annual national conferences on weights and measures and the *Journal of Weights and Measures* was set up in 1908 for the "benefit of Dealers, Sealers, and the Purchasing Public." During World War I, the War Industries Board pushed producers to standardize products and conserve resources, so much so that colors of typewriter ribbons were reduced from 150 to 6 colors, automobile tires from 287 types to 9, and buggy wheels from 232 sizes to 4 (Cochrane 1966). After the war, the American Standards Association was formed in 1919 and persuaded hundreds of firms to adopt common standards with the active support of the National Bureau of Standards and the then-commerce secretary, Herbert Hoover, in a bid to improve efficiency.

Large private corporations such as General Motors, General Electric,

Westinghouse, and American Telephone and Telegraph also set up special purpose departments to establish standards for devices. American Telephone and Telegraph was reputed to have saved a million dollars because of the use of a standardized repeater in long distance lines (Chase and Schlink 1927, p. 235).

Consumer leagues.—Unlike standards and testing organizations that were extensions of professions and trade associations, consumer leagues were loosely modeled on trade unions. If trade unions used the threat of a strike to discipline employers, consumer leagues relied on the threat of a boycott to push for better prices and standards of living. Generally, these consumer leagues had visible involvement from women activists and, on occasion, trade union leaders and often depended on women's groups for support.

The earliest such organization, the National Consumer League, was founded in 1899 to ameliorate the plight of child labor and had many women members. Florence Kelley, the first executive director, a former chief factory inspector and attorney, stated that the major aim of the organization was to "moralize" the demand function of consumer power so that workers received fair living wages, goods were produced under sanitary conditions, and the interests of the community were promoted (Kelley 1899, p. 290). By contrast, other consumer leagues were concerned about the effect of prices on the standard of living. Bureaus of labor statistics in different states had legitimized the linkage between wages, prices, and consumption by conducting cost-of-living inquiries. Massachusetts was the first to create a state labor bureau of statistics, and by 1885, 14 other states had followed suit. By 1910, cost-of-living inquiries referred to the consumer as a category of actors (Samson 1980, p. 114).

Rising food and meat prices motivated attempts to organize consumer leagues. The National Anti-Food Trust League started in 1909, had the president of District of Columbia's Central Labor Union as a member (Samson 1980, p. 152), but it failed to go beyond organizing a few hearings. A consumer boycott of meat in 1910 was organized in Cleveland by the city's central labor union, and 100,000 people signed a pledge not to eat meat for 30 days. Soon, boycotts sprang up in Pittsburgh, Kansas City, and Baltimore, but they failed because of poor organization, the lack of solidarity, and the inability of meat producers to bear lower prices.

By 1910, a few buying clubs consisting of housewives had also begun to appear on the horizon. The same year the Housekeepers Cooperative Association was organized in Pittsburgh with the intent of providing advice on spending and on defeating unscrupulous practices. The Federated Marketing Club was organized in Chicago in 1911 to publish daily lists of bargains and had a board consisting of the presidents of the Illinois Federation of Women's Clubs, the Cook County Federation of Women's

Clubs, and the Chicago Federation of Labor. In 1911, a Housewives League was started in New York to ensure accurate weights and measures. A butter boycott was successfully prosecuted, the National Association of Housewives' Leagues was established in 1912, and a magazine called *Housewives League* was rolled out to "represent the consumer in the fight to reduce the cost of living" (Samson 1980, p. 160). The First World War saw the league championing conservation, but by the end of the war, consumer leagues had become moribund due to the lack of support and the waning of inflation.

Distinctiveness of nonprofit CWOs.—Nonprofit CWOs were an important departure from standard-setting organizations and consumer leagues. Unlike standard-setting bodies that were extensions of trade associations and professional societies, nonprofit CWOs were formally separate from the industries they monitored and were explicitly oriented to the needs and interests of consumers. Unlike consumer leagues that were extensions of women's clubs or relied primarily on women for support, nonprofit CWOs served both men and women. If consumer leagues sought to ensure lower prices, nonprofit CWOs aimed to protect consumers (by deterring fraud, promoting product safety), to educate consumers (by providing literacy and decision-making skills), to lobby legislators, and to inform consumers (via best buy recommendations in the consumer goods, health, and consumer finance industries). How was this form established? How did it emerge as a departure from its organizational precursors? How were its boundaries defined? Below, I address these issues by discussing how a resource space or niche initially arose for nonprofit CWOs.

Emergence of a Resource Space for Nonprofit CWOs

Rising expenditures on consumer durables, mounting complexity of product choices, changes in the pattern of advertising, and lack of product liability rules created the social context that made it possible for institutional entrepreneurs to mobilize resources for the establishment of nonprofit CWOs as solutions.

Durable goods expenditures.—Disposable income, in 1970 dollars, registered a fivefold increase from \$14.1 billion in 1901 to \$74.5 billion in 1930 (U.S. Bureau of the Census 1975). As disposable income grew, individuals were able to purchase a wider array of consumer durables, and they began to account for a greater share of disposable income (Olney 1991). Table 1 shows changes in the average allocation of durable goods expenditures. It reveals that the share of transport vehicles jumped from 6.4% in the period 1899–1908 to 28.3% in the period 1919–28. The spurt rise in automobile purchases was spurred by installment buying and by the growth of personal finance companies that allowed individuals to pur-

TABLE 1

AVERAGE ALLOCATION OF DURABLE GOODS EXPENDITURES (%), 1900-1928

	1899-1908	1909-18	1919-28
Transport vehicles	6.4	10.3	28.3
Furniture	30.8	24.8	22.5
Household appliances	3.3	3.7	4.1
Entertainment appliances	1.0	1.3	1.5
All major durable goods	41.6	40.0	56.5
China and tableware	19.5	23.0	14.0
House furnishings	21.1	18.5	14.7
Watches and jewelry	6.9	7.8	5.5
Medical products	1.0	1.9	1.9
Books and maps	7.9	6.6	5.6
Miscellaneous	2.0	2.2	1.9
All minor durable goods	58.4	60.0	43.5
Total	100	100	100

SOURCE —Olney (1991), pp. 37-45

NOTE —All estimates are based on constant 1982 price estimates using Bureau of Economic Analysis data.

chase goods through credit. Furniture and household appliances accounted for virtually all other major consumer durable goods spending. Electrification provided an impetus for the purchase of household appliances. In 1907, only 8% of homes in America had electricity, but by 1925, 53.2% had electricity (U.S. Bureau of the Census 1975). Households with electricity were a lucrative market for clothes washers, refrigerators, and radios. Old products like the broom, unchanged since ancient Egypt, were superseded by new products like the vacuum cleaner (Silber 1983, p. 10). Table 1 also shows that china, tableware, watches, and jewelry were the most purchased minor consumer durable goods.

As products like the icebox were replaced by automatic electric refrigerators, consumers had to choose technologically complex products with which they had little familiarity. Moreover, individuals with wealth, as Veblen (1919) noted, were able to construct personal styles through the purchase of consumer durables. These styles were imitated by others in the social ladder so that the status conferred by a consumer durable good mattered more than its usefulness. Indeed, products enabled individuals to extend their definition of self and to augment their personal identities as distinctive individuals (Schudson 1978).

Complexity of product choice.—Product comparison became costlier with an increased number of price classes and a rising number of brands within each price class. Table 2 shows that consumers could buy cars in four price classes and often had to choose among price categories within

TABLE 2
AUTOMOBILE SALES BY PRICE CLASSES (%), 1903-26

Car Price (in dollars)	1903	1908	1913	1918	1923	1926
Low:						
< 675	4.2	5.0	46.6	41.6	62.2	51.6
Medium:						
676-875	22.1	18.8	.5	11.6	7.5	8.5
876-1,375	42.4	19.5	19.9	33.5	13.2	24.7
High						
1,376-1,775	6.1	10.6	13.7	6.0	8.8	7.5
1,776-2,275	1.7	10.4	10.2	3.9	5.2	3.2
2,276-2,775	12.0	16.6	3.9	1.1	1.4	2.4
Luxury:						
2,776-3,775	7.6	3.5	2.6	1.7	1.2	1.3
3,776-4,775	2.6	10.4	1.6	.4	.4	.5
> 4,776	1.3	5.2	1.0	.2	.1	.3

SOURCE.—The data on percentage of shares by share classes are drawn from Epstein's (1929) survey of auto producers.

a price class. In each price class, a purchase involved a comparison of several producers. In 1918, for example, the medium-priced producers, who accounted for 43.2% of all cars sold, included Willys-Overland, Buick, Dodge, Chevrolet, Maxwell, Studebaker, Saxon, Reo, Regal, Briscoe, Grant, Hupp, Dort, and Mitchell, each of whom allowed consumers to choose a plethora of options (Thomas 1977).

By 1926, annual model changes powered by the doctrine of planned obsolescence led to an emphasis on style and aesthetic appeal rather than mechanical performance. Paid advertising for automobiles exploded from \$1 million a year in 1901 to \$4 million in 1914 and \$41 million in 1926 (1.5% of the wholesale value of industry output; Thomas 1977, p. 236). Thus, consumer choice began to be influenced by style rather than substance.

The changing pattern of advertising.—In 1900, \$542 million was spent on advertising through newspapers, radio, outdoor, and miscellaneous media. By 1930, advertising expenditures had grown to \$2.6 billion (U.S. Bureau of the Census 1975). Newspapers, especially Sunday newspapers, began to cover fashion, etiquette, beauty, hygiene, and clothing by having special "advice" columns for women on how to use products to acquire social status and to signal it discreetly (Schudson 1978, pp. 100-102). In 1902, *Good Housekeeping*, a Hearst publication, developed a "Seal of Approval" for goods it judged worthy of advertising space in its pages.

Table 3, drawn from Pollay (1985) documents changes in the pattern

TABLE 3
CHANGING PATTERNS OF ADVERTISING, 1900–1930

DIMENSION	% OF ADS EXAMINED			
	1900s	1910s	1920s	1930s
Advertisers				
Manufacturers	48	80	82	94
Direct marketers	50	17	12	4
All other	2	2	5	2
Rhetorical focus:				
Source (ethos)	11	7	14	24
Tone (pathos)	27	25	30	42
Logic (logos)	62	68	56	35
Depiction:				
Product only	19	21	14	9
People only	29	15	23	17
Product and people	34	57	61	72
Neither	18	7	2	3
Tactical focus:				
Product features	59	43	29	20
Benefits to be gained	40	54	62	63
Risks to be avoided	1	4	8	17

SOURCE — Pollay (1985)

of advertising during the period 1900–1930.⁸ It reveals changes in who advertised: The share of direct marketers dropped from 50% in the 1900s to 4% in the 1930s whereas advertisements from manufacturers rose from 48% in the 1900s to 94% in the 1930s. The rhetorical focus of advertising also changed. Advertisements based on appeals to logic shrank from 62% in the 1900s to 35% in the 1930s, and advertisements appealing to emotion (pathos) rose from 27% in the 1900s to 42% in the 1930s. Advertisers sought to manipulate greed, envy, and fear. Listerine evoked the fear of halitosis in its campaign, “even your best friend won’t tell you,” and in concert with other products, these campaigns led to the discovery of bodily functions as targets of marketing opportunity (Fox 1984). The content of the advertisement also changed as advertisements depicting products and people shot up from 34% in the 1900s to 72% in the 1930s. Celebrity testimonials extolled products: the queen of Romania explained why she entrusted her skin to Pond’s cold cream in advertisements, and movie stars like Joan Crawford shared why they used Lux soap. Finally, table

⁸ Pollay’s (1985) analysis is based on a random sample of 2,000 advertisements spanning the period 1900–1980, which were culled from the 10 leading magazines during that time period.

TABLE 4
PRODUCT LIABILITY CASES IN AMERICAN COURTS,
1900-1930

Decennial	No of Product Liability Cases	No of Negligence/ Tort Cases Cited
1900-1910	3	59
1911-20	4	28
1921-30	7	14

NOTE.—Product liability cases are from Madden (1988) and Vandell (1989). Negligence and tort cases cited in the product liability cases come from court reporters (e.g., *Northeast Reporter*, *Southeast Reporter*, *Southwest Reporter*, *Federal Reporter*, *Pacific Reporter*).

3 shows that the tactical focus of advertisements was also transformed. In the 1900s, 59% of advertisements emphasized product features, but by the 1930s, only 20% did so. Instead, the 1930s advertisements began to stress the benefits accruing from the product and the risks that could be avoided by using the product. Taken together, all of these changes suggest that, as Veblen (1919) lamented, the fabrication of consumers had become a routine operation much like a mechanical industry.

The lack of product liability rules.—A 1923 study of 244 advertisements by Stuart Chase found 44 advertisements that were palpably false and 28 cases where harmful products were touted as useful (Vaile 1927). But consumers possessed restricted legal rights and had little recourse when faced with defective or harmful products of poor quality. The period 1860-1920 was dominated by the doctrine of negligence, wherein ideas of privity, contributory negligence, and assumption of risk were applied to deny plaintiffs recovery and to impede consumer rights (Spacone 1985). Under the notion of privity, manufacturers had no obligation to a party to which it was not in direct contract (e.g., passengers in a car could not sue the manufacturer for a defect in the car). Similarly, the assumption of risk meant that consumers had accepted some risk when they purchased a product or used a service. Due to these restrictive ideas, there were very few product liability cases, and the few cases that were decided invariably cited negligence and tort cases to legitimate their claims. Table 4 provides a decennial count for the period 1900-1930. It shows that the number of product liability cases was in single digits and that, as the number of cases inched upward, the tendency to cite negligence and tort cases to justify claims of liability diminished. One landmark case, *MacPherson v. Buick Motor Co.*, was decided in 1916 by the New York Court of Appeals, and Justice Cardozo weakened the privity requirement when he ruled that a manufacturer was liable if any product was inherently dangerous or if it

was dangerously and negligently manufactured. However, the small number of product liability cases suggests that the idea of liability was not firmly established before 1930. Legal historians suggest that it was only from 1930 onward that scholars, especially Friedrich Kessler and Fleming James, led the movement for tort reform by attacking the premises of contract and negligence law and by inventing the idea of enterprise liability (Priest 1984). It was only in the late 1950s and early 1960s that the *MacPherson v. Buick Motor Co.* case was followed in several jurisdictions and served to undermine the ideas of privity and negligence (Spacone 1985, p. 22).

CWOs as Impartial Testing Agencies: Consumers Research

Sporadic attempts backed by women's organizations to improve the lot of the consumer arose but did not flourish. Thus, attempts by the Federated Marketing Club in Chicago (1911), the Housewives' League in New York (1911), and the National Association of Housewives' Leagues (1912) to provide bulletins of prices failed as inflation declined. It was against this backdrop that two issue entrepreneurs, Stuart Chase and Frederick Schlink, diagnosed the problems facing consumers and framed a new social control mechanism as the solution—the CWO.

The impartial testing frame.—Chase, an accountant by profession who was affiliated with Veblen in Howard Scott's ephemeral Technocratic movement, wrote two polemics entitled *The Challenge of Waste* (1922) and *The Tragedy of Waste* (1925) and castigated the riot of products that were detrimental to man and outside the category of wants. With Frederick Schlink, who was an ex-employee of the National Bureau of Standards and the National Standards Association, Chase in 1927 wrote *Your Money's Worth*, which became an instant best-seller. Chase and Schlink portrayed the consumer as an "Alice" in a wonderland created by advertising and product differentiation. They blamed manufacturers for failing to *serve* the consumer by instead creating wasteful variety. They noted that a "housewife needing a sewing machine needle might as well look for one in a haymow as on a neighbor's machine. Such needles are made in nine diameters . . . and in lengths varying by as little as one thirty-second of an inch" (Chase and Schlink 1927, p. 174). Chase and Schlink (1927, p. 2) attacked advertising for deceiving consumers, deriding the "conflicting claims, bright promises, fancy packages, soaring words, and almost impenetrable ignorance" and asserted that modern marketing practices impeded consumers' access to the benefits of mass production.

Consumers were urged to imitate Schlink's consumer's club established in a church in White Plains, New York. This neighborhood club prepared two "confidential lists"—one carrying "products considered to be of good

value in relation to their price; the second, products, one might well avoid, whether on account of inferior quality, unreasonable price, or of false and misleading advertising" (Chase and Schlink 1927, p. 254). The book sparked hundreds of inquiries, and Schlink transformed the neighborhood club into Consumers Research (CR)—an organization that sought to serve as an "economist, scientist, and an accountant" (Silber 1983, p. 18). The list was renamed as the *Consumers Research Bulletin* that would "investigate, test, and report reliably hundreds of commodities" (Silber 1983, p. 18).

The organization embodied a new mechanism for the social control of business (its target) and consumers (its clients). It explicitly stated that its aim was to readjust the "rapidly increasing power of the manufacturer and seller over the mind and judgment of the consumer" (CR 1932, p. 1). As a social control device, CR aimed to promote new norms of consumption. Appendix table A1 delineates the norms propagated by CR.

The new organization sought to reform the system of production so that waste was reduced and producers *served* customers by making goods needed by customers and charging fair prices. The tools of control were product standards and scientific tests. Efficiency was a valued norm and was to be promoted by reducing wasteful variety through standardization. CR also aimed to replace magic in salesmanship with the norm of scientific analysis, so that facts rather than emotions would guide consumption. Consumers were exhorted to be rational and vigilant and to use the information provided by tests to extract the most value for their dollar. Finally, the organization sought to signal its impartiality through its nonprofit structure and its distance from political parties.

Legitimizing the Impartial Testing Frame

As an embodiment of a new form of social control, CR encountered very little opposition except some stray attacks, despite its vigorous assaults on manufacturers and advertisers. One writer foresaw a future society in which a man who did not like standard coffee was arrested for buying an unauthorized brand (Stoddard 1928, pp. 117–20). But in the main, CR's attacks evoked little opposition from business firms, advertisers, or even state authorities.

The absence of conflict does not mean that CR's birth was a plain story of an organization emerging to fulfill a need in the marketplace. Instead, it suggests that CR's founders had deftly defused opposition by their creative use of hitherto unconnected cultural elements. Granovetter (1994, p. 484) notes that the assembly of elements into a form is a "good example of the Schumpeterian definition of entrepreneurship, which involves pull-

ing together previously unconnected elements." Biggart (1989) shows that the synthesis of unconnected elements can disarm opposition to innovations—hence, direct-selling organizations seamlessly linked the logics of the family and marketplace and were able to legitimate themselves. Put simply, skillful framing allows entrepreneurs to set the context and to exercise unobtrusive power. Ranson, Hinings, and Greenwood (1980, p. 8) note that "power is most effective and insidious in all its consequences when issues do not arise at all."

CR's founders preempted opposition by seamlessly combining hitherto unrelated cultural elements into a new frame or model of consumption. CR's founders deftly framed their critiques of business and advertising around the ideas of service to the customer and truth in advertising—concepts that businessmen and advertisers had begun to implement in a bid to professionalize their trades. CR's founders also borrowed elements of their solution—standards, testing, and science—from the work of industrial standard-setting bodies and the home economics profession. Below, I suggest that these professional infrastructures shaped the cognitions of CR's founders but shielded them from opposition. By borrowing ideas from these cultural repositories and by aligning themselves with professionalization efforts, CR's founders were also able to insulate themselves from potential attacks.

The professionalization of retailing.—Between 1910 and 1925, numerous business firms adopted codes of ethics that exalted service to the customer (Heermance 1924). A study by Kitson (1922) in the *Journal of Political Economy* reported that advertisements that contained the word "service" increased from less than 5% in 1900 to 24% in 1920. The growth of the service ideal was due to the professionalization of business and competitive pressures on firms.

Schools of retailing blossomed across the country between 1902 and 1920. In 1902, Arthur Sheldon set up a school of scientific salesmanship in Chicago, and by 1921, 20 such schools existed (Samson 1980, p. 204). The American Association of Collegiate Schools of Business was created in 1916, and by 1925, it had 25 members, many of whom instituted courses in retailing and business ethics built around ideas of service. The service idea diffused among businesses also because of the Rotary Clubs. Initially set up in 1905 as a mutual-aid society among individuals from different commercial fields, the Rotarians championed the service ideal after the 1911 convention when their platform, influenced by Arthur Sheldon, stated that "the science of business is the science of service; he profits most who serves best" (*National Rotarian*, March 1912, p. 15). By 1915, the Rotary Club adopted a code of ethics that made service the "Golden Rule," and it became the model for business firms.

The promotion of the service ideal by the retailing profession induced

business firms, especially retailers, to implement the service idea as a competitive weapon. A survey by the U.S. Bureau of Education in 1917 found that 50 stores, such as Wanamaker's, Altman's, Macy's, and Grant's, employed teachers of salesmanship. Stores began to rely on one-price selling and money-back guarantees and even employed "service shoppers" to make reports on the quality of service. Store manuals on service such as Wanamaker's exhorted employees to "hold the customer's interest jointly sacred with the interests of business" (Fuld 1916, p. 16). Trade associations such as the National Retail Dry Goods Association spoke of the need for the retailer to serve as a "protector of the consumer" (Letts 1914, p. 7) and the *Journal of Retailing* echoed this sentiment. Schlink published an article in the journal a month before his best-seller with Chase appeared and urged retailers to work as guardians of consumers and to reduce the costs of shopping. It was CR, however, that came to take on the role of a guardian of consumers, devoted to ensuring that business served customers.

The exaltation of the service ideal provided an evaluative yardstick for Chase and Schlink and made it possible for them to expose contradictions between precepts and practices. When Chase and Schlink (1927, p. 12) attacked business for violating norms of service and lamented that "we are deluged with things we don't wear, which we lose, which go out of style, which make unwelcome presents for our friends," businessmen and retailers could not respond with counterattacks because to do so would have meant disowning the ideal of service.

Professionalization of advertising.—Similarly, when Chase and Schlink (1927, p. 12) castigated "the advertiser [who] plays on the essential monkey within us," the advertising community could not respond because its bid to professionalize itself hinged on truth in advertising. After the formation of the Sphinx Club of New York in 1896, numerous advertising clubs sprang up, and in 1904, several clubs merged to form the Associated Advertising Clubs of America (AACA). By 1909, the AACA had established an educational committee to systematize training and to promote honest advertising. In the 1911 convention of the AACA, a proposal initiated by John Romer, the publisher of *Printers Ink*, was introduced wherein it was illegal for an advertisement to contain deceptions or misleading facts. Romer also urged advertising clubs to establish vigilance committees to ascertain the truthfulness of advertisements issued by members, to discipline errant members, or even to take them to court. The National Vigilance Committee was created in 1912, and by 1914, 24 cities had also founded vigilance committees in a bid to signal the advertising community's commitment to probity and professional conduct. In 1916, vigilance committees were renamed as Better Business Bureaus, and by 1930, more than 10,000 businesses were supporting these bureaus in numerous cities (Samson 1980, p. 343). Some of the bureaus instituted programs akin to

service shopping, wherein advertising claims were tested by purchasing the products. However, these bureaus lacked any enforcement capability—for example, a resolution banning the use of paid testimonials was rarely honored. Similarly, in 1925 the New York bureau found that the claim made by Macy's that its prices were at least 6% less than other department stores was unjustified. Macy's resigned from the bureau because it did not accept the finding. The Macy's case undermined the credibility of the bureaus and showed that they could not contend with large and influential firms.

Thus, the advertising community sensitized CR's founders to the issues of ethics and truth and laid the foundation for their critique. When CR's founders attacked advertising for deception and characterized it as "largely competitive wrangling as to the relative merits of two undistinguished and often indistinguishable compounds" (Chase 1925, p. 125), advertisers could neither organize any vigorous defense nor launch an offensive. By co-opting elements of the truth-in-advertising movement into their critique of advertising, CR's founders insulated themselves from attack.

Home economics profession and rational purchasing.—A new profession, that of home economics, had emerged around 1908, and its propagation of rational techniques of purchasing provided invaluable protective cover for Chase and Schlink. Ellen Richards, a lecturer in sanitary chemistry at MIT, sought to eradicate inefficiency in home management by organizing conferences of like-minded women at Lake Placid, New York, during the period 1899–1908. These conferences persuaded librarians to put books related to home management under the "economics of consumption" category rather than the category of "production." In 1908, the American Home Economics Association was created to foster thrift, efficiency, and safety in the "management" of households.

Home economists suggested that poverty also was the result of incompetent spending. In 1910, a number of charitable societies, especially the Cleveland Associated Charities, established separate home economics departments and integrated consumer expertise into social work practice (Zelizer 1994, p. 153). Home economists sought to substitute carelessness and ignorance in home management with systematic budgeting, comparison shopping, and rational techniques to allocate resources to clothes, rent, and food. The emphasis on rational purchasing also induced advertisers to change their techniques; some advertisers began to emphasize paired product comparisons and automobile manufacturers, such as Buick, whose cars won races, built their advertising around the theme "tests tell" (Schudson 1978).

Colleges began to introduce home economics courses in their curricula due to funding from the federal government through the Bureau of Home

Economics and Cooperative Extension Services of the Department of Agriculture. The latter organized a large project under which 2,500 home economists involved with numerous land-grant colleges set up demonstration projects—each such project covered 12–35 communities in an area and brought women together once a month to discuss consumption (Sorenson 1941, p. 64). As the number of students enrolled in federally aided home economics programs rose from 31,000 in 1918 to 175,000 in 1930 (U.S. Bureau of the Census 1975), techniques of rational home management and purchasing also spread. Additional support for the home economists' attempt to rationalize purchasing by households received encouragement from the U.S. Commerce Department's National Bureau of Standards, which published pamphlets for housewives. Two pamphlets called *Measures for the Household* and *Materials for the Household* published in 1915 and 1917 urged consumers to rely on standards and tests and contributed to the rise of a new model of purchasing by consumers.

Growth of CR.—The emphasis on testing and standardization disarmed opposition to CR and fueled its growth. Pfeffer (1992, pp. 247–49) says, "Power is most effectively employed when it is fairly unobtrusive. Using rational, or seemingly rational processes of analysis helps render the use of power and influence less obvious. . . . Decisions are perceived to be better and are accepted more readily to the extent they are made following prescribed and legitimate procedures."

CR grew quickly. In 1927, it had 656 subscribers, but by 1933, there were 42,000 subscribers (Silber 1983). The lack of sustained opposition from advertisers and businesses and Chase and Schlink's skillful framing of their cause and solution spurred CR's growth. CR's ideas grew more popular due to the Depression, which forced consumers to be more vigilant about price-quality relationships. For example, as shirts, sheets, and shoes wore out quicker than consumers expected them to, people felt cheated and wanted to remedy the situation (Sorenson 1941, p. 10). A wave of exposés also fostered interest in product testing. *Skin Deep* by Mary C. Phillips, *The American Chamber of Horrors* by Ruth DeForest Lamb, and *100,000,000 Guinea Pigs: Dangers in Everyday Foods, Drugs, and Cosmetics* by Kallett and Schlink, the secretary of CR, called for mandatory disclosure of product ingredients because weak laws rendered "a hundred million Americans . . . as unwitting test animals in a gigantic experiment with poisons, conducted by food, drug and cosmetic manufacturers" (cited in Silber 1983, p. 18). A spate of other books followed, and much like the muckraking journalism, this guinea-pig journalism blended moral indignation with insistent pleas to the public to use objective information in purchasing. As a result, by 1935, CR had a staff of 50 employees and 200 outside consultants, and the readership of the *Consumers Research Bulletin* rose to 55,000.

CWOs as Engines of Radical Change: Consumers Union

As CR grew, other consumer advocates were lured to the organization by its rising stature. Notable recruits included Alexander Crosby (a journalist critical of advertising in newspapers), Dewey Palmer (an engineer seeking relief from college teaching), and Arthur Kallett (a colleague of Schlink's at the American Standards Association). As new activists joined, older activists left CR because of disagreements with Schlink. In 1932, Stuart Chase left the board of CR, and in 1933, two other board members, Bernard Reis and E. J. Lever, resigned due to budgetary disagreements with Schlink. These members were replaced by newcomers such as J. B. Matthews and Mary C. Phillips, who were allied with Schlink.⁹

These budgetary disagreements within CR stemmed from differences about how much money should be spent on testing and how much on advocating reform of labor conditions. Some employees and members of the board, Palmer and Kallett, for example, felt that CR should lobby for the reform of labor conditions. The 1932 edition of the *Consumers Research Bulletin* (p. 3) noted that "CR is desirous of adding union and other labor groups to its subscription list; second, members of CR's board are pretty generally interested in not ignoring the labor conditions under which products are produced." By contrast, Schlink, Matthews, and Phillips thought that social questions concerning wages and working conditions were beyond their scope and could not be studied in an "unbiased" manner (Sorenson 1941, p. 47). Thus, CR was beset with a debate about whether it would be an impartial provider of information for consumers or a journal of radical political economy dedicated to the amelioration of the plight of workers.

These tensions flared during a union recognition drive by CR's employees in August 1935. When 20 or more employees formed a chapter of the Technical, Editorial, and Office Assistants Union and asked for recognition, John Heasty, a chemist and union president, and two other union activists were fired. A strike was called, and the union twice demanded the removal of Mary C. Phillips and J. B. Matthews from the board because they were concerned with "personal aggrandizement" and not with the "real interests of the masses of consumer-workers at heart" (Strike Demands of Workers, cited in Katz 1977, p. 75). Dewey Palmer was dismissed from the board because he opposed Schlink, and Kallett threw in his lot with the striking workers. Frederick Schlink, Mary Phillips, J. B. Matthews, and other board members resorted to strikebreakers, legal injunctions, and armed detectives to retain control of the magazine and the organization. The Association of Subscribers, composed of members

⁹ The five-member board was dominated by this triumvirate.

sympathetic to labor concerns, was formed to exhort Schlink and his allies to negotiate with the union but had little effect, and as a result, the association demanded a refund of subscriptions.

Attempts to break the deadlock between the management and the strikers were fruitless. An attempt by Reinhold Niebuhr to mediate was rejected by Schlink, who also challenged the jurisdiction of the National Labor Relations Board when it was called in to settle the strike. Schlink and his allies felt that the strikers were "dupes of business" and "Communists" (see Silber 1983, p. 21). As a result, 30 workers led by Kallett established a new organization called Consumers Union (CU), which sought to unite the cause of consumers and workers. Initially, three names were considered—National Consumers Union, Consumers Technical Union, and Consumers Union of the United States. The first was discarded because it was similar to the National Consumer League (which focused on child labor). The use of the word "technical" in the second name created opposition. Finally, the third was chosen as the best description of the new organization. Its publication was to be called *Consumers Union Reports*.

The radical change frame.—CU was more than a breakaway faction of CR and was premised on a strikingly different diagnosis of the problems facing the consumer. The organization's founders felt that a watchdog guarding consumers could not merely provide scientific and impartial information to consumers to make rational purchasing decisions. Instead, the originators of CU defined the consumer as a worker concerned with the standard of living and not just a rational actor seeking to get the best value for his money. The problem facing consumers was not one of variety and deceptive advertising, it was also one of wages and income. The founders opined that "all the technical information in the world will not give enough food or enough clothes to the textile workers' families living on \$11 a week. . . . They like the college professor or the skilled mechanic are ultimate consumers. . . . The only way in which any organization can aid them materially is by helping them, in their struggle as workers, to get an honest wage" (*Consumers Union Reports*, May 1936, p. 2).

CU's founders proposed that the new organization "give information and assistance on all matters relating to the expenditure of earnings and family income . . . to create and maintain decent living standards for ultimate consumers." CU was more than an advocacy organization and aspired to be a new method for the social control of business (its target) and consumers (its clients). Appendix table A2 details the norms promoted by CU. The new organization viewed socially responsible buying and collective action as tools to improve working conditions and create decent living standards for consumers. Like its rival, CR, CU also valued the norms of scientific analysis and rational purchasing and was committed to impartial testing. In contrast to CR, however, CU saw buying as a socially responsi-

ble act, urged members to picket antiunion stores, and pleaded with them to use labor conditions as a criterion in the purchasing process. CU was also committed to the norm of equity in incomes and saw itself as a critic of companies exploiting their workers. Its founders were careful to signal their commitment to impartial testing by not only having a nonprofit organization but also by decoupling product ratings from evaluations of labor conditions.

The first issue explained why *Consumers Union Reports* was a confidential magazine (because of legal reasons), outlined its rating system (why products were rated as “best buys,” “also acceptable,” and “not acceptable”), ran an exposé of Alka Seltzer’s advertising claims, and compared credit unions and commercial banks. By the end of 1936, CU issued a 240-page buying guide that also discussed labor conditions in firms making consumer products. Subsequent issues analyzed products and labor conditions: for example, *Consumers Union Reports* used blind tests to assess major cigarette brands, found no significant differences in taste, and concluded that it made sense to select the cheapest union-made cigarette. It opined that the big four cigarette manufacturers had labor policies that “ranged from bad to worse. . . . [Their] social policy is transplanted from the Dark Ages” (*Consumers Union Reports*, August 1938, p. 19). In 1936, 11 of its members were arrested picketing Orbachs, the department store in New York. CU supported antifascist boycotts, provided testimony in Federal trade-law hearings, organized promotional drives among the poor, and supported labor movement picketing during 1936–38. It was so involved that, of the 50 employees of CU, only seven were technical employees (Silber 1983, p. 27). CU’s board reflected the endorsements of the trade unions and left-leaning intellectuals. It had several radical trade unionists among its officers, and the board was composed of academics (Robert Brady, Charles Marlies), union leaders (Philip Randolph, Rose Schneiderman, Heywood Broun), an accountant, and the staff representative (John Heasty).

Contested Boundaries of the Nonprofit CWO Form

The rival frames promoted by the founders of CR and CU led to divergent definitions of the boundaries of the nonprofit CWO form. CR was focused in its goals and technology and how it served clients: it sought to reduce waste, rely on testing, and catered to one constituency—consumers. By contrast, CU was diversified in its scope—its goals, technologies, and constituencies. It sought to provide information to consumers, to improve standards of living, to rely on testing and boycotts, and to serve two constituencies—consumers and workers

These contradictory proposals suggested that the support base of the

nonprofit CWO form was heterogeneous. Social movement researchers suggest that a heterogeneous potential support base permits a range of definitions of the situation (Zald and McCarthy 1980). If CR drew its support from those committed to professionalizing retailing and advertising and rationalizing home management, CU drew strength from those committed to labor issues. During the period preceding the birth of CU, labor legislation had created a propitious environment for trade unions; in 1932, the Norris-La Guardia Act gave workers the right to join unions, and in 1935, the Wagner Act recognized the right of employees to strike and defined unfair labor practices on the part of employers. Consequently, union membership rose from 1.1 million in 1901 to 3.75 million in 1935. As unions became more militant, work stoppages increased from less than 2,000 to 4,500 in 1937 (U.S. Bureau of the Census 1975). Hence, linking labor and consumer issues was both topical and sensible. For a new organization that sought to espouse a radical agenda, CU grew rapidly—by the end of 1936, it had 20,700 members, and by 1937, it had close to 40,000 members.

Moreover, neither form enjoyed clear-cut technical advantages premised on transaction cost considerations. CR's claim to efficiency hinged on the premise that focus (on a single goal, technology, and constituency) reduced coordination costs. This argument implies that focus on a single issue credibly communicates commitment to a cause (Williamson 1985), reduces conflicts of interest among donors, and precludes adroit managers from exploiting conflicts among donors and from diverting resources toward causes they personally favor (Fama and Jensen 1983). By contrast, CU's claim to efficiency pivoted on the premise that diversification leads to economies of scope in lobbying. When market segments are similar, goals and technologies are related to each other and can augment each other, then, managers can use a dominant management logic and realize efficiency gains (Aldrich et al. 1994).¹⁰

Since neither definition of form enjoyed any decisive technical advantages, the choice of which frame to use to design nonprofit CWOs and

¹⁰ The question of whether focus is more efficient than diversification is an empirical one. Research on the for-profit sector shows that related diversification is more efficient than unrelated diversification, and the evidence that indicates that single-product firms outperform diversified firms is hampered by a selection bias—such firms are typically long-established firms with invisible assets, and it may be a case of capabilities constraining diversification, rather than focus leading to capabilities (Oster 1994, p. 193). In the nonprofit sector, Aldrich et al. (1994) found that diversified trade associations did not have higher death rates than specialist trade associations that were focused on one industry. Edwards and Marullo's (1995) analysis of peace movement organizations suggests that national-level organizations were more viable if they pursued broader agendas.

what boundaries to impose on their activities hinged on politics and, in particular, endorsements by powerful actors—professions, state authorities, and naturally, the targets of watchdog activity—businesses and advertisers.

Early Supporters of Impartial Testing and Radical Change Frames

At the outset, CU constituted only one instance of the model of a watchdog pursuing labor and consumer agendas, and CR was the only instance of an organization incarnating the model of the CWO as an impartial testing agency. However, the scales began to tilt against CU when a new organization, Intermountain Consumers Service, was set up by S. A. Mahood, a chemist affiliated with the U.S. Products Lab and a consumer educator to boot. It had a membership of 3,500 subscribers, and it chose to mimic CR and extol the model of the consumer as a rational decision maker who needed to rely on scientific tests.

The model of the consumer as a rational decision maker, promoted originally by CR, was also diffusing through governmental agencies and professional societies. Thus, the Consumers Advisory Board was established in the National Recovery Administration in 1933, and it sponsored a report by Robert Lynd, the sociologist, which called for the creation of common standards and the allocation of funds for testing. The Consumer Counsel was established in the Department of Agriculture in 1933 that issued a bimonthly publication called *Consumer Guide* with a circulation of over 150,000. Later, when the National Emergency Council set up a consumer's division in 1934 to coordinate all consumer-related activities of the federal government, 200 county consumer councils were established to facilitate two-way communication (Sorenson 1941, p. 16–17).

The concepts of rational decision making, standardization, and scientific testing that were initially promoted by CR also spread in professional circles. The American Home Economics Association set up the Standing Committee on the Standardization of Consumer Goods in 1927, became a member of the American Standards Association in 1928, and organized cooperative standardization projects between women and merchants in several cities (Sorenson 1941, p. 66). Later, other long-established national organizations, such as the National Education Association, the American Federation of Teachers, the American Marketing Association, the Mid-Western Economics Association, and others began to devote time to consumer education issues during their annual conferences.

By contrast, small newly founded consumer groups such as the Consumer Conference in Cincinnati (1934), the League of Women Shoppers (1935), the High Cost of Living Conferences (1935), and the Milk Consumers Protective Committee (1939) endorsed CU. The dearth of endorse-

ments only invigorated the desire of CU's founders to increase circulation as a means of jump-starting a bandwagon of support and pushing for radical reform. Resources from increased circulation were essential to fund an expansion of the product-testing program and to reward technical talent. Moreover, CU was wary about creating local groups because it did not want to be legally responsible for their conduct. Additionally, CU's founders had also realized that their bulletin was not reaching low-income workers and felt that boosting circulation could help them to sustain low-price editions of their bulletin.

Opposition to the Radical Change Frame

CU's attempt to increase circulation evoked resistance from diverse institutional actors. For example, in 1937, CU published "The Contraceptive Report," which summarized the efficacy of birth control techniques—it chose this topic given the belief that family planning was important for sustaining standards of living for workers. It was also a shrewd bid to sell more copies of *Consumers Union Reports*. However, the postmaster general of New York banned it in 1939, and CU fought a court case that was only resolved in its favor in 1943. CU sought to advertise itself through mass media but faced a media boycott. Sixty-two newspapers, including the *New York Times*, refused to sell advertising space to CU (Sorenson 1941, p. 37). The *Times* stated that the reason for the refusal was that CU's "services comprised attacks on industries. This the *Times* does not permit" (letter by C. McPuckett to S. Zaslow, cited in Katz 1977, p. 170). Attempts to advertise in professional journals such as *Science* met with the reply that "it is not possible for the Consumers Union to supply information that has scientific validity" (letter by J. McCatrell to A. L. Kallett, 1938, cited in Katz 1977, p. 171). Even the *Journal of Home Economics* declined to accept advertisements because they felt that CU was making accusations that could not be substantiated, and home economists rejected invitations to join the board.

Rival entrepreneurs also constructed CWOs that were interpreted as threats by CU's founders. In 1937, Albert Lane began to publish *Consumer Bureau Reports*, which provided favorable ratings in return for free samples from manufacturers. Undeniably, this venture threatened to besmirch the image of all CWOs. However, CU was more concerned than CR because Lane's publication had a name that could be confused with CU's own *Consumers Union Reports*. Moreover, Lane's publication also copied CU's breezy format and may have confused readers because it praised products that CU had rated as unacceptable. Another entrepreneur, William Foster, established the Consumers Foundation as a research organization advising consumers, mobilized money, and even succeeded

in drafting Robert Lynd as a supporter. CU raised questions about the new organization's sources of funds, and it was found that the Institute of Distribution, a trade association of chain stores, had supported it. Attempts such as these tarnished the image of all CWOs, but CU was concerned about spillover effects.

CU was also attacked by other publications, especially the Hearst newspapers. In its second issue, CU asked its readers to support a strike against Hearst's *Wisconsin News*, and from 1936 to 1939, it issued articles that exposed Hearst's Good Housekeeping Institute (so much so, that the Federal Trade Commission launched an investigation against the institute). As a result, Hearst became a bitter enemy and charged that CU was a Communist-front organization. The *Women's Home Companion* also accused CU of undermining the American way of life.

The attacks reached their zenith in 1938 when a House committee on subversive activities chaired by Congressman Dies sought to investigate whether CU was engaged in un-American activities harmful to the national interest. J. B. Matthews, an associate of Schlink's at CR, served as counsel for the Select Committee on Un-American Activities (the Dies Committee) and Matthews suggested that Kallett's writings and the fact that a CU ex-employee, Susan Jenkins, had admitted to being an employee of a Communist newspaper (the *Daily Worker*) were proof that the organization was a Communist front. Matthews labeled CU as a "red transmission belt"—the Hearst newspapers printed Matthews's accusations in full, and *Good Housekeeping* distributed the Matthews report in the November-December issue of its Consumer Information Service. However, the Dies Committee did not challenge the credibility of CU's rival, CR.

One reason why CR did not face problems from the Dies Committee is that Matthews was an associate of Schlink. Another reason is that the very existence of CU may have made CR more acceptable. Haines (1984) suggests that "radical" organizations by their sheer existence make "moderate" organizations more acceptable and enable them to acquire support and resources. It is possible that CR may have been shielded from criticism from business corporations and newspapers.

Narrowing Boundaries: Adaptation of CU

CU sued Albert Lane not only because he was deemed to be a copycat using a similar name and style but also because suing Lane may have positioned CU as a defender of not just its own integrity but the standing of the nonprofit CWO form. Rising to the defense of a form was a credible method of signaling CU's commitment to independence and impartiality at a time when its standing was being impugned by CR and the business press. CU won the suit. It secured some protection from the damage done

by the Dies Committee when Eleanor Roosevelt rushed to its defense in a variety of public statements. The Dies Committee was running out of money, and some members of the committee, especially Representative Voorhis, attacked Chairman Dies for the conduct of the hearings. The Dies Committee was a select committee, which later, in 1945, was made a standing committee—the House Un-American Activities Committee. Although there was no systematic investigation of the charges leveled at CU, the fact that CU's name was on the committee's shortlist of targeted Communist organizations was a matter of concern to the founders and potential supporters. Later, these charges would be revived during the McCarthy era when CU's founders were forced to testify.

Nevertheless, the pressure exerted by hostile activists (such as Matthews), elements of the media (the Hearst publications), and politicians (such as Dies and others on the Select Committee on Un-American Activities), impelled CU's founders to disengage from radical advocacy. Kallett and some other officers of CU recognized that the organization would be shielded from external attacks if it developed its testing program and concentrated on rating products. The adaptation of CU was a gradual process of disengaging from the labor agenda.

During the second annual meeting in 1938, three resolutions urging a focus on ratings and a disavowal of interest in labor and the threat of fascism or other "ideologies" were introduced but not approved. A founder and board member, Charles Marlies, tendered his resignation (Sorenson 1941, p. 49). In 1938, CU decided to "lapse" its coverage of labor conditions in its buying guide because "detailed information and numerous qualifications were usually necessary to give a fair picture of labor relations in a given industry or plant, labor notes are not included. In addition, the very swift changes continually taking place in labor relations make it impossible to include in annual publication, annual data that would remain up to date and reliable" (CU 1938, p. 9).

CU's management rationalized their retreat by noting that a truce between labor and business was essential in combating the specter of fascism in Europe. CU's management justified product testing for its "trickle-down effects"—the work of the testers would improve the lot of the workers indirectly by enhancing the quality of all merchandise and by promoting legislation benefiting all consumers. In a turnabout in 1939, CU began to assert that "just the ordinary products bought each day can save members \$50 to \$300 a year" (*Consumers Union Reports*, April 1939, p. 14).

In 1939, Arthur Kallett and Colston Warne, the president, sought to derive support from the scientific community and arranged a meeting with the Cambridge-Boston chapter of the American Association of Scientific Workers (AASW), a fledgling organization that was created in 1938 by a

group of biologists at Woods Hole, Massachusetts. At this meeting, CU members confessed to the inadequacy of their testing, and the AASW agreed to provide expert advice and testing for certain products. In 1940, the board rejected an attempt by the editor of *Consumers Union Reports* to recruit union members as subscribers. In 1940, the National Advisory Committee, composed of academics, was also created to establish linkages with colleges and universities.

World War II forced CU to adapt to a war economy, where shortages rather than abundance and inflation rather than quality became the problems. Product testing made sense in an economy of abundance but not in an economy of rationing. In 1939, CU began to publish a section on war and prices that was expanded in 1940 into a weekly newsletter called *Bread and Butter*. This publication distanced itself from labor issues but sought to assist low-income consumers by acting as an early warning system designed to help consumers cope with shortages and to buy goods likely to be hit with price increases by profiteers. *Bread and Butter* exhorted consumers to unite against profiteers—the Council of Organized Consumers was established in 1942 but was unsuccessful. The war economy, however, forced CU to shrink its payroll and the length of its bulletins and took a toll on membership.

In 1942, CU sought to invite Robert Lynd to join their board, but he demurred even as he denied that he was a “red-baiter” (Katz 1977, p. 201). The legacy of the Dies Committee’s allegations convinced Kallett and his colleagues that they needed to more formally distance themselves from a labor agenda. In 1942, the name of the magazine was changed from *Consumers Union Reports* to *Consumer Reports* to ostensibly avoid confusion in the mind of the public about whether a company’s labor policies influenced the ratings of its products.

Kallett and his colleagues were concerned about the tendency of some of CU’s employees to espouse radical agendas through *Bread and Butter*. Even at the time of the formation of the newsletter, there were divisions in the board between whether to forewarn consumers of price increases or to educate them about product maintenance. As resource constraints impelled Kallett to downgrade the paper quality of *Bread and Butter*, its advocates within CU accused Kallett of curtailing resources for the newsletter. In 1944, Kallett persuaded the board that *Bread and Butter* and *Consumer Reports* ought to be delinked, and as a result, consumers were not to be forced to take both publications simultaneously. In 1946, with the end of the war, *Bread and Butter* outlined a six-point postwar reconstruction proposal that included price controls, rationing, food subsidy programs, rent control, subsidized housing, and health care. At this time, the business community was lobbying Congress for an end to ra-

tioning and price controls. Influenced by the legacy of the Dies Committee, Kallett was concerned about ideological conflict and the advocacy of *Bread and Butter* and in 1947, persuaded the board to collapse *Bread and Butter* into a feature of *Consumer Reports* under the title "Economics for Consumers." In June 1948, Leland Gordon, an economist at CU, asked the organization to reinstate reports on labor conditions, but Kallett and Isserman (a prolabor sympathizer) objected on the grounds that it might compromise CU.

The containment of advocacy proceeded within CU. In 1947, Kallett changed CU's bylaws so that subscribers could be distinguished from members and thereby sought to deflect criticism that members might be Communists. In the same year, the union of CU's workers lamented that its board representative was a second-class citizen and asked that the position be terminated. The board agreed with alacrity. This was a radical turnabout for an organization that had prided itself on being a collectivist enterprise at the time of its creation and was, for the first year or so, structured as an egalitarian organization.

The postwar years also saw the rise of Keynesian economics, and some of CU's leaders felt that a marriage of consumer reform goals with the Keynesian philosophy of full employment and the importance of consumption would best situate CU in a regulated free-enterprise economy. Indeed, as early as 1943, some of the more radical elements at CU, working with *Bread and Butter*, wanted to work for the establishment of a consumer democracy, but Kallett and others concerned more with CU, felt that accepting Keynes's ideas meant that the organization would not be labeled un-American.

Hence, CU's conversion was not one dramatic event but a slow and gradual process of repositioning and reconstruction. CU slowly ceased to be an engine of political, social, or moral activism and reinvented itself as an impartial testing agency. Over time, CU began to recognize products as conditionally acceptable even if they were not best buys and expanded its testing to ensure that it was done in-house instead of subcontracting it out as CR did. Its evaluations and options appeared in the form of tables and charts with numerical results of tests. Annual surveys of its members enabled it to respond to its subscribers' needs. For example, CU provided ratings of cars on the basis of speed after it learned that consumers valued speed as a criterion in purchasing cars. Even after introducing a segment on health and medicine in 1945 and providing careful summaries of medical research on smoking, *Consumer Reports* provided information on how to "roll your own" cigarettes during the cigarette famine of 1945.

By emphasizing testing and science and disavowing radical labor advocacy, Kallett and his colleagues not only shielded CU from external attack

but transformed it into a scientific conservative to fit with the prevailing beliefs about science, rigor, and objectivity. The price of viability for the organization was abstinence from advocacy.

CR may have won the struggle over what a nonprofit watchdog ought to be, but it was outdistanced in the battle for circulation. By joining the media and politicians in critiquing CU, CR may have sown the seeds for its own decline because its embrace of scientific conservatism proved advantageous to CU. By 1949, with 500,000 subscribers, CU rated 1,793 brands spanning 116 products, and its technical division was divided into electronics, textiles, automobiles, special projects, chemistry, and foods. By contrast, CR had made little progress and refused to publicize its circulation details (Thorelli and Thorelli 1974). CU's gradual adaptation to political pressure and its concurrent adjustment to the demands of a post-war economy fueled a rise in CU's membership base. Figure 1 chronicles CU's growth from its inception to 1971. After an initial rise from 1936 to 1938, the circulation of *Consumers Union Reports* decreased due to World War II and slowly picked up after 1945. After the war, CU received the endorsement of prestigious universities and the federal government. The Home Economics Department of Cornell University assisted in the testing of sewing machines and the Department of Agriculture allowed an employee to work under its supervision on the grading of canned food.

In 1951, the Dies Committee's allegations were revived by Better Business Bureaus in Akron, Cincinnati, Dayton, and Detroit, and the school boards in these cities threatened a ban on CU's magazines in the school libraries. CU avoided publicity on the bannings to the extent possible and denied allegations directly when asked. Its restraint stemmed from concern because its name was still on the list of un-American organizations compiled by the House Un-American Activities Committee. J. B. Matthews joined McCarthy's investigative staff in 1953 and resurrected the charges. A Hearst columnist, E. F. Tompkins, publicized them, and CU's board stated that "in the present period of ideological conflict, the Board of Directors of Consumers Union reaffirms its faith in democratic society in which the production of goods and services is guided by the free choice of consumers" (CU 1952, pp. 1-2). CU's leaders testified before the House Un-American Activities Committee, and after a long hearing, the organization's name was finally struck off the list of un-American organizations in 1954.

By 1955, William Whyte noted that *Consumer Reports* and *Consumers Research Bulletin* had an impact over and above their actual circulation—both magazines were produced by organizations that institutionalized distrust of advertising and aimed to create impersonal trust by providing scientific, objective, and impartial evaluations. By 1956, the ratings

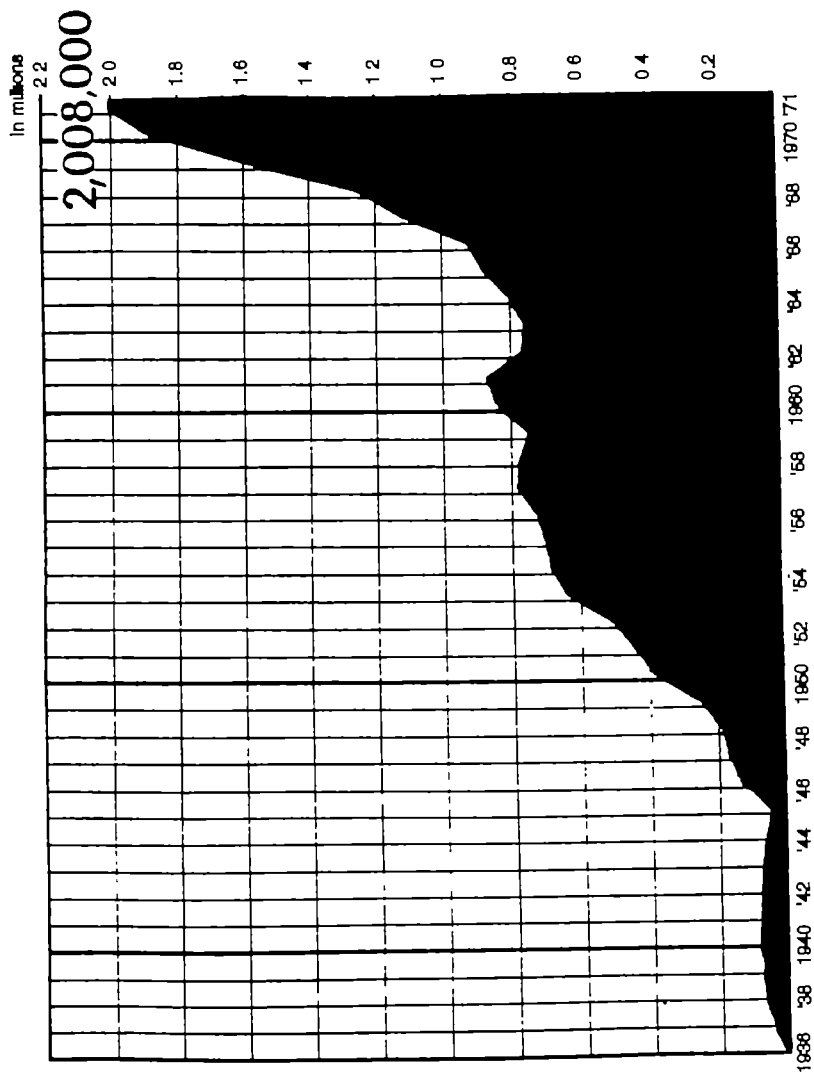


FIG 1.—*Consumers Union Reports* and *Consumer Reports* circulation, 1936–71

of CU could spell good fortune or doom for a company: spokesmen for Maytag washers and Volkswagen cars attributed their penetration of hitherto oligopolistic markets to favorable ratings (Samson 1980). In 1959, the *New York Times*, an erstwhile critic, noted that "what Dr. Spock is to freshman parents, the Consumers Union is to bewildered housewives" (Silber 1983, pp. 30–31). By 1958, psychologists and marketing experts opined that consulting the scientific ratings provided by CWOs had become a buying ritual in an age of science that enabled consumers to think of themselves as rational, logical, and informed consumers (Sargent 1958). So established was the identity of CWOs as meticulous scientific conservatives that the ratings techniques used by them were lampooned by magazines. A British comedy group even had a skit entitled "*Consumer Reports* looks at religion."

However, the transformation of CU into a scientific conservative was a double-edged sword. If it strengthened the legitimacy of CU, it also inhibited it from espousing radical causes and thereby provided an opening for other issue entrepreneurs to articulate fresh grievances and create new nonprofit CWOs. Ralph Nader's (1965) *Unsafe at Any Speed*, an exposé of General Motors, was initially offered by him to CU, but the book was disparaged by its automotive division as unscientific and biased. Nader called CU a sleeping giant, and later his best-seller enabled the formation of new groups dedicated to the enforcement of the legal rights of consumers.¹¹

DISCUSSION

This study responds to Scott's (1995, p. 147) call to study the emergence of new organizational forms and to illuminate the coevolution of culture and organizations. One implication of this study is that organizational forms are structural conveyances for cultural materials that become infused with norms, values, and beliefs as entrepreneurs use frames to mobilize resources. When entrepreneurs use frames to legitimate the new form, they inject an organizational form with cultural content and serve as conduits by which cultural rules are encoded into organizations. The infusion of prevalent but unconnected cultural elements into the watchdog form was not an apolitical act. Rather it exemplified the use of unobtrusive power, wherein problems of consumption and solutions were framed in much the same way that a picture is framed (Pfeffer 1992). CR's founders, Schlink and Chase, deftly co-opted the ideas of service and truth in adver-

¹¹ Later, Nader was inducted into the board of CU, and the organization sought to support other consumer groups through grants

tising in their critique of business for wasteful inefficiency and advertising for deception and preempted opposition. Similarly, the prevalence of industrial standards and home economics training enabled Chase and Schlink to present their solution—an independent testing agency designed to aid rational decision making by consumers. In sharp contrast, CU's founders, Kallett and his colleagues, strove to frame their organization as a champion of downtrodden workers and as a crusader for decent and just living standards. Thus, as both groups of founders relied on the Western cultural account with its themes of progress and justice (Meyer, Boli, and Thomas 1987), they generated conflict about the practical implications of that cultural account. By showing how entrepreneurs are skillful political operatives employing frames as mobilization devices (Swidler 1986; Biggart 1989; Snow and Benford 1992), this article addresses DiMaggio and Powell's (1991, p. 30) lament that neoinstitutional accounts of new organizational forms say little about "how skillful entrepreneurs put multiple institutional logics to strategic use" and creates a bridge between social movement theory and the literature on the creation of new forms (Fligstein 1996a).

The case study of the origins of nonprofit CWOs illustrates how the boundaries of organizational forms are defined by institutional processes rather than by transaction cost considerations. Organizational economists insist that transaction cost considerations determine the "efficient boundaries" of organizational forms (Williamson 1985). However, Powell (1991) notes that institutional processes shape the boundaries of forms when technical differences among rival forms are ambiguous or when the criteria to evaluate technical excellence are disputed. Institutional processes, as Hannan and Freeman (1989) point out, can erode or establish boundaries. Bricolage by entrepreneurs, copying errors, and deinstitutionalization can efface boundaries among forms. By contrast, closure of social networks, collective action, endorsement by powerful actors, and taken-for-grantedness can secure the boundaries of organizational forms.

Although complexity of product choice, deceptive advertising, and the lack of product liability rules created the resource space for nonprofit CWOs and made it possible for quality assurance activities to be bundled into a new organizational form, the boundaries of the form emerged as an outcome of ideological competition. CR's founders, Schlink and Case, borrowed the idea of scientific testing from standard-setting bodies and the idea of reducing the cost of product comparisons from discussions in the retailing profession. They sought to erode the boundaries between nonprofit CWOs and standard-setting organizations. By contrast, CU's founders, Kallett and his allies, strove to import characteristics of trade unions into the consumption sector and sought to make nonprofit CWOs the counterparts of unions. These rival combinatorial attempts repre-

sented two potential ways of bundling the activities of the nonprofit CWO form. Neither model enjoyed a decisive technical advantage—if CR's focus reduced its coordination costs, then CU's diversified emphasis led to economies of scope in lobbying.

Each model was premised on a different notion of the identity of the consumer. Chase and Schlink defined a consumer as a decision maker keen on getting the best value for the money, promoted norms of efficiency, rationality, and scientific analysis, and extolled watchdogs as impartial testers. By contrast, Kallett and his allies viewed the consumer as a worker keen to better his or her standard of living, promoted norms of socially responsible buying and equity, and portrayed watchdogs as engines of radical change. CR and CU, therefore, represented alternative models for the social control of industry, premised on different ideals of identity.

Sustained hostility from the Hearst publications, especially *Good Housekeeping*, and politicians such as Congressman Dies forced CU to scale back labor-related activism and emphasize scientific testing. Thus, the dikes of vested interest prevented the rise of the consumer-as-worker model and instead confined consumers and CWOs into a narrower channel of activity. By emphasizing how a contest over the identity of the consumer delineated the boundaries of the form, this article reveals how consumers and consumption were produced by processes other than industrialization (see DiMaggio 1994), shows how struggles over identity (White 1992) are the mechanisms by which forms are imbued with meaning and constituted as cultural objects, and speaks to DiMaggio's (1991) call to incorporate agency and interest in the creation of new forms.

Another implication of this article is that, just as routines become operative only when there is a cease-fire agreement among organizational members (Nelson and Winter 1982), new organizational forms become operative only when they embody a truce among contending social factions. Had conflict over who a consumer was and what consumption was persisted, given the precarious interpersonal relationships between the founders of CR and CU, the energies of both organizations would have been sapped by internecine warfare. Conceivably, an opening may have been created for business firms, newspapers, and advertisers to discredit the notion of nonprofit CWOs. When CU's founders bowed to pressure by embracing the model of an impartial tester and disavowed sociopolitical advocacy, there was a cessation of hostilities and a moratorium on the debate about the identity of consumers and nonprofit CWOs. This truce on the contours of the nonprofit CWO form was maintained because CU's founders were keen to avoid the risks of political pressure from Congress and the defensive alertness of actors opposed to sociopolitical advocacy. The absence of debate about the role of nonprofit CWOs made it possible

for them to rationalize consumption and to become influential monitors of big business. Although nonprofit CWOs embodied an implicit truce, truces can also be explicit and encoded in legislation. For example, there was intense debate about the scope of commercial banks, insurance firms, investment banks, and securities firms before the passage of the Glass-Steagall Act of 1933; the enactment of this law codified a truce among the proponents of the interests of these constituencies. Explicit or implicit truces can be breached by new issue entrepreneurs who strive to take advantage of new technologies and the emergence of new constituencies with grievances. Thus, Nader breached the truce on nonprofit CWOs by exploiting the issue of product safety and contributed to the rise of nonprofit CWOs as legal activists.

The study of nonprofit CWOs also speaks to DiMaggio's (1991) call to shed light on the role of professionals in the birth of new forms. Attempts to professionalize occupations delegitimize the existing social order and create openings for institutional entrepreneurs to build new forms. For example, museum directors launched a reform movement that discredited existing museums and spawned new organizational models. In the case of nonprofit CWOs, attempts by sales experts to train retailers in customer service sought to professionalize retailing but delegitimized the existing system of managing customer relationships. Similarly, the truth-in-advertising movement aimed to professionalize advertising but discredited prevalent modes of advertising. Thus, both movements provided an opening for Chase and Schlink to launch their critique of business and advertising and to legitimate nonprofit CWOs as desirable and appropriate social control mechanisms.

The saga of nonprofit CWOs also illuminates how the state can be an arena for disputes about the boundaries of an organizational form. Just as Congress and the courts can be audiences for jurisdictional disputes between professions (Abbott 1990), they can also serve as audiences for jurisdictional contests between rival coalitions of entrepreneurs. The debate between the faction that saw CWOs as engines of radical change and the opposing faction that preferred CWOs to serve as impartial testers spilled into Congress when J. B. Matthews, a critic of the radical change model (and CU), and a one-time board member at CR, exploited his role as a counsel to the Dies Committee in investigating CU. These investigations tainted the reputation of CU, led to its inclusion in a roster of un-American organizations, and induced CU to scale back advocacy. The specter of Congressional investigation continued until the McCarthy hearings, forced CU to embrace the model of the watchdog as impartial tester, and maintained the truce about the contours of nonprofit CWOs.

The story of CU also demonstrates that mimicry can be an outcome of

coercion. Neoinstitutionalist accounts treat the state and the professions as engines of coercive and normative isomorphism and portray mimicry as the unthinking adoption of peer group practices (DiMaggio and Powell 1983). Neoinstitutional accounts of mimicry paint cultural change as a process of diffuse social learning and neglect how conflict underlies the assimilation of new models into the general culture (see Tarrow 1989). Our case study suggests that vested interests can foist coercive cultural expectations on an organization and can create inducements for it to imitate its peers. March and Olsen (1989) suggest that conflicts may be resolved through the logic of aggregation and give and take or the logic of integration, wherein one of the parties can learn from the other and even convert to the other's point of view. In our case study, the logic of integration prevailed—CU responded to withering political attacks by discarding its founding beliefs, by embracing the ascendant logic, and by converting itself into a scientific conservative.

A related implication that flows from the conversion of CU is that resistance may be followed by acquiescence. Neoinstitutionalist researchers generally describe acquiescence, compromise, avoidance, defiance, and manipulation as alternative strategic responses to institutional pressures (Oliver 1991). However, these responses may also be sequential and shaped by strategic considerations. CU began as an insurrection against the prevailing frame, and when its attempt to resist failed, its founders complied to external pressure and copied the ascendant model of nonprofit CWOs. An alternative was to concentrate on a narrow subscriber base. The founders of CU chose to acquiesce because they were concerned about the survival of the organization and its long-term prospects.

Some limitations of this study also deserve elaboration and point to future research possibilities. This study analyzed a phase of the consumer movement characterized by the existence of two ideological frames and two main organizational embodiments. One frame became ascendant because it enjoyed greater political support. The organization backing the losing frame embraced the ascendant frame, thereby making it unquestionably dominant. However, when multiple frames and multiple organizational embodiments exist, the process by which one frame becomes dominant may be more path dependent (Powell 1991) or may be shaped by density-dependent processes of legitimation and competition (Hannan and Freeman 1989).

As a historical analysis, this study only looked at the origins of nonprofit CWOs during their formative phases in the United States but not at their transformation by Ralph Nader and environmental activists. Tarrow (1989) suggests that collective action unfolds in cycles of protest and urges that the idea of structural stability needs to be replaced with the notion

of dynamic stability—the successive realizations in different forms of the same principle. Cycles of protest are characterized by heightened conflict across the social system, diffuse in specific ways from the center to the periphery, possess peaks and troughs, and are the crucibles from which new social repertoires evolve. Since their origins, CWOs have been profoundly influenced by cycles of protest—for example, the Civil Rights and antiwar movements provided an ideological context for Nader and other issue entrepreneurs to undermine the model of CWOs as scientific conservatives and to transform them into legal activists seeking to enforce the legal rights of consumers. Later, the environmental movement also led to an expanded notion of corporate social responsibility and the ethics of consumption. Clearly, the question of how protest cycles underpin the structural transformation of nonprofit CWOs merits scrutiny. One payoff may well be a more nuanced account of how social movement industries evolve and how interorganizational relationships within a social movement industry are transformed through time (Zald 1992).

This study lacked a comparative thrust and confined its attention to the United States. Nonprofit CWOs have diffused across the world since the establishment of the first CWOs in the United States (1927 and 1935), Germany (1925), and Finland (1939). In 1995, Consumers International, the international association of nonprofit CWOs, had 200 members from 90 countries (Consumers International Bulletin 1997). The international evidence reveals considerable diversity in the origins of nonprofit CWOs. In France, Belgium, the Netherlands, and Austria, for instance, CWOs were affiliated with labor organizations. By contrast, CWOs were created as distinct entities but were affiliated with women's guilds in Finland, Canada, Denmark, and Switzerland. Except the Swiss CWO, other CWOs evolved to focus exclusively on consumer issues and roots in women's guilds atrophied. In Norway, product testing began in 1939 with support from state authorities, and now several state-sponsored CWOs flourish in Asia and Africa—many of these are government affiliate members of Consumers International. The birth of nonprofit CWOs in other developed and newly industrializing countries needs to be chronicled in order to understand the international trajectory of this social control mechanism.

Additionally, the relationships between nonprofit watchdogs and for-profit watchdogs also merit scrutiny. Although this study did discuss the antagonism between CU and *Good Housekeeping* during the early history of nonprofit CWOs, detailed analyses of the interdependencies between nonprofit CWOs and for-profit counterparts are needed. When do nonprofit CWOs spawn for-profit rivals? When do nonprofit CWO publications such as *Consumer Reports* and for-profit organizations such as J. D.

Powers play complementary roles? Do issue entrepreneurs bear the risks of legitimating new social control arrangements only to see forprofits engage in successful free riding? The systematic analysis of such questions is needed to not only extend our grasp of CWOs but also to promote a better understanding of the division of labor between the for-profit and the nonprofit sector.

This study focused on the process by which nonprofit CWOs came to be injected with norms and beliefs. It is also necessary to analyze how these social control mechanisms influence culture. Just as cultural theories spur structural innovation, social structures also contribute to the rationalization of society. As noted earlier, nonprofit CWOs embodying the model of scientific conservatives played a part in rationalizing the purchasing decisions of individuals, making them logical and systematic (Sargent 1958). Future research needs to delineate how environmental watchdogs have influenced conceptions of environmental risk and have rationalized the management of environmental disasters. Similarly, the analysis of how human rights watchdogs have influenced conceptions of rights and have formalized oversight of nation-states also deserves scrutiny.

Moreover, this article has confined its attention to one type of nonprofit policing organization—the nonprofit CWO. Other nonprofit watchdogs also clamor for attention. The analysis of human rights watchdogs (epitomized by Amnesty International) and watchdogs that monitor corruption in nation-states (such as Transparency) might shed light on the social processes fostering the rise of control mechanisms that issue report cards on modern states. An irony of all watchdog organizations is that, like other social control mechanisms that institutionalize distrust, they are also vulnerable to the abuse of trust (Shapiro 1987). For example, the Clean Water Act permits environmental watchdogs to act as a private attorney general and to sue polluters who break the law. Many such suits are settled before they are brought to court; polluters agree to halt the flow of effluents and to make charitable donations to environmental groups not involved in the suit, ostensibly, to repair the damage done by the polluter. Although a judge authorizes such consent decrees, environmental watchdogs have been known to divert resources to satellite organizations, which may pursue projects unrelated to the consent decree. Even when contested by the Justice Department, such settlements between an environmental watchdog and its targets have been upheld by courts as consistent with the provisions of the Clean Water Act (Felten 1991). Who polices watchdogs? Do failures of watchdogs promote the growth of new social control mechanisms? Is the price of impersonal trust an iron cage, as forecasted by Weber?

APPENDIX

TABLE A1
NORMS PROMOTED BY CR

Norm and Mechanism	Rationale
Efficiency:	
Supplying standards	<p>The variety of goods "should be kept within finite and rational limits." (Chase and Schlunk 1927, p. 9)</p> <p>Buying "an electric toaster entirely on attractiveness . . . which if looked at unemotionally must be admitted to be unrationalizable, wasteful, and . . . incapable of . . . giving the result that must be our ultimate purpose—namely, the production of goods of merit at a cost that is reasonable." (Schlunk 1927, p. 181)</p> <p>"A standard is based flatly on the assumption that an economic process has no justification other than that of supplying things which mankind needs." (Chase 1921, p. 284)</p>
Scientific analysis:	
Testing by CR	<p>CR's purpose is to "conduct, promote, and generally encourage scientific technical investigations and research into the nature of commodities and services in use by consumers and of interest to them." (CR 1932, article 3)</p> <p>"If science could displace magic in salesmanship, the whole curve of consumption would change." (Chase and Schlunk 1927, p. 163)</p>
Rationality.	
Consumers' use of tests	<p>Consumers will achieve "better results by learning how to get more from their money than by continually striving to get more income." (<i>CR Annual Cumulative Bulletin</i>, 1940, p. 2)</p> <p>CR "emphasizes function rather than style. It works sincerely in the interests of consumers as it believes consumers ought to be independent and rational and getting the most physical value for their money" (Ayres 1934, p. 162)</p> <p>"The consumer will be ill-advised if he expects any organization to carry out for him the functions which as a consumer he himself should and must fulfill." (Chase and Schlunk 1927, pp. 264–67)</p>

TABLE A1 (*Continued*)

Norm and Mechanism	Rationale
Impartiality	
Nonprofit structure and distance from political parties	The aim is to be a "clearing house for consumers' information, . . . to cooperate voluntarily, and without gain or profit with any individuals, firms, government agencies, associations and corporations which are engaged in kindred pursuits." (CR 1932, article 3)
	CR "holds no brief for any particular kind of economic conduct or social or political order." (Chase and Schlunk 1927, p. 3)

NOTE.—Sources of control are standardization and consumer preferences based on tests. Objects of control are to eliminate waste and to ensure that producers serve customers

TABLE A2
NORMS PROMOTED BY CU

Norm and Mechanism	Rationale
Scientific analysis:	
Testing by CU	CU's purpose is to "supervise and conduct research and tests" to provide information for the consumer. (CU 1936)
	Scientific advice is a "far better guide to intelligent purchasing than any other ordinarily available to the consumer." (CU 1938, p. 6)
Rationality:	
Consumers' use of tests	The use of tests by the consumer "substitutes the best technical knowledge for haphazard guessing; and it gives the consumer the satisfaction of having his buying choices determined by technical tests than by the cleverness of an advertising copy writer." (CU 1938, p. 6)
Independence:	
Decoupling quality ratings from labor conditions	"Labor conditions under which many products are manufactured are described in the monthly <i>Reports</i> , although these conditions do not affect the quality and price ratings." (CU 1938, p. 6)

TABLE A2 (Continued)

Norm and Mechanism	Rationale
Social responsibility.	
Purchasing power	<p>"Members ought to help as consumers in such struggles by not buying from anti-union stores " (Editorial after the picketing of Orbach's department store in New York, <i>Consumers Union Reports</i>, July 1937, p. 24)</p> <p>"Ratings which are accompanied by labor notes when they originally appeared in the <i>Reports</i> carry a reference line. . . Members are urged to consult these notes for general labor information, but should remember they are specifically applicable to conditions prevailing at the time of writing " (CU 1938, p. 9)</p> <p>Fighting for "higher wages is not enough" and workers must "fight for fair prices and good quality in the products they buy " (<i>Consumers Union Reports</i>, May 1936, p. 24)</p> <p>"These profiteers with their price rises, shortages and sabotage are helping Hitler. Fighting them means a more effective fight for victory." ("Selling <i>Bread and Butter</i>," press release, CU 1942)</p> <p>The wartime consumer is "improving his effectiveness as a citizen from the wealth of news and information relating to his interest in every issue of the <i>Reports</i> and <i>Bread and Butter</i> " (<i>Consumer Reports</i>, September 1942, p. 226)</p>
Equity.	
Action and criticism	<p>CU's purpose is to "aid individuals and group efforts of whatever nature and description in seeking to obtain, create and maintain decent living standards for ultimate consumers." (CU 1936)</p> <p>All the "technical information in the world will not give enough food or enough clothes to the textile worker's family living on \$11 a week " (<i>Consumers Union Reports</i>, May 1936, p. 2)</p> <p>"The Electrical Appliance Industry has the distinction of enjoying the full measure of recovery and at the same time indefensibly exploiting its employees " (<i>Consumers Union Reports</i>, July 1936, p. 9)</p> <p>"The sit-down is nothing but the extension of the picket line to the machines. It is the workers answer to the denial of civil rights on the picket line." (<i>Consumers Union Reports</i>, January-February 1937, p. 27)</p>

NOTE.—Sources of control are consumers' preferences based on tests and responsible buying and collective action. Objects of control are to sustain decent living standards for ultimate consumers and to reform working conditions.

REFERENCES

- Abbott, A. 1990. *The System of the Professions*. Chicago: University of Chicago Press.
- Aldrich, Howard E. 1979. *Organisations and Environments*. Englewood Cliffs, N.J.: Prentice-Hall.
- Aldrich, Howard E., Catherine R. Zimmer, Udo H. Staber, and John J. Beggs. 1994. "Minimalism, Mutualism, and Maturity. The Evolution of American Trade Associations in the 20th Century." Pp. 223-39 in *Evolutionary Dynamics of Organisations*, edited by J. Baum and J. V. Singh. New York: Oxford University Press.
- Astley, W. G. 1985. "The Two Ecologies: Population and Community Perspectives on Organizational Evolution." *Administrative Science Quarterly* 30:224-41.
- Baron, J. N., F. R. Dobbin, and P. D. Jennings. 1986. "War and Peace: The Evolution of Modern Personnel Administration in U.S. Industry." *American Journal of Sociology* 92:350-83.
- Becker, Howard. 1963. *Outsiders: Studies in the Sociology of Deviance*. New York: Free Press.
- Biggart, Nicole W. 1989. *Charismatic Capitalism: Direct Selling Organisations in America*. Chicago: University of Chicago Press.
- Briloff, A. J. 1972. *Unaccountable Accounting*. New York: Harper & Row.
- Brint, Steven, and Jerome Karabel. 1991. "Institutional Origins and Transformations. The Case of American Community Colleges." Pp. 337-60 in *The New Institutionalism in Organizational Analysis*, edited by W. W. Powell and Paul DiMaggio. Chicago: University of Chicago Press.
- Brittain, W. J., and John Freeman. 1982. "Organizational Proliferation and Density Dependent Selection." Pp. 291-338 in *The Organizational Life Cycle*, edited by J. R. Kimberly and R. H. Miles. San Francisco: Jossey-Bass.
- Brobeck, Stephen. 1990. *The Modern Consumer Movement: References and Resources*. Boston: G. K. Hall.
- Bykerk, L., and A. Manney. 1995. *U.S. Consumer Interest Groups: Institutional Profiles*. Greenwich, Conn.: Greenwood.
- Campbell, John L., and Leon Lindberg. 1990. "Property Rights and the Organization of Economic Activity by the State." *American Sociological Review* 55:634-47.
- Carroll, Glenn R. 1984. "Organizational Ecology." *Annual Review of Sociology* 10: 71-93.
- Carroll, Glenn R., Jacques Delacroix, and Jerry Goodstein. 1990. "The Political Environment of Organizations: An Ecological View." Pp. 67-100 in *Evolution and Adaptation of Organisations*, edited by Barry Staw and Larry L. Cummings. Greenwich, Conn.: JAI Press.
- Chase, Stuart. 1922. *The Challenge of Waste*. New York: New York League for Industrial Democracy.
- . 1925. *The Tragedy of Waste*. New York: Macmillan.
- Chase, Stuart, and Frederick Schlink. 1927. *Your Money's Worth*. New York: Macmillan.
- Cochrane, Rexmond C. 1966. *Measures for Progress: A History of the National Bureau of Standards*. Washington, D.C.: Department of Commerce.
- Consumers International. 1997. *Consumers International Fact Book*. New York: Consumers International.
- CR (Consumers Research) 1932. "Bylaws." New York: Consumers Research.
- CU (Consumers Union). 1936. "Certificate of Incorporation." New York: Consumers Union.
- . 1938. *Consumers Union Buying Guide*. New York: Consumers Union.
- . 1952. "Minutes of the Board of Directors." New York.
- DiMaggio, Paul. 1988. "Interest and Agency in Institutional Theory." Pp. 3-21 in

- Institutional Patterns and Organizations: Culture and Environment*, edited by L. G. Zucker. Cambridge, Mass.: Ballinger.
- . 1991. "Constructing an Organizational Field as a Professional Project: U.S. Art Museums 1920–1940." Pp. 267–92 in *The New Institutionalism in Organizational Analysis*, edited by W. W. Powell and Paul DiMaggio. Chicago: University of Chicago Press.
- . 1994. "Culture and Economy." Pp. 27–57 in *Handbook of Economic Sociology*, edited by N. J. Smelser and R. Swedberg. New York: Russell Sage.
- DiMaggio, Paul J., and Walter W. Powell. 1983. "The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields." *American Sociological Review* 48:147–60.
- . 1991. Introduction to *The New Institutionalism in Organizational Analysis*, edited by Walter W. Powell and Paul DiMaggio. Chicago: University of Chicago Press.
- Douglas, Mary. 1986. *How Institutions Think*. Syracuse, N.Y.: Syracuse University Press.
- Durkheim, Émile. 1950. *The Rules of Sociological Method*. Glencoe, Ill.: Free Press.
- Edwards, Bob, and Sam Marullo. 1995. "Organizational Mortality in a Declining Social Movement: The Demise of Peace Movement Organizations in the End of the Cold War Era." *American Sociological Review* 60 (6): 908–27.
- Epstein, Ralph C. 1928. *The Automobile Industry: Its Economic and Commercial Development*. Chicago: A. W. Shaw.
- Fama, Eugene, and Michael Jensen. 1983. "Separation of Ownership and Control." *Journal of Law and Economics* 24:301–23.
- Felten, Eric. 1991. "Money from Pollution Suits Can Follow a Winding Course." *Insight* 7:18–20.
- Fligstein, Neil. 1996a. "How to Make a Market: Reflections on the Attempt to Create a Single Market in the European Union." *American Journal of Sociology* 102:1–33.
- . 1996b. "Markets as Politics: A Political Cultural Approach to Market Institutions." *American Sociological Review* 61:656–73.
- Fornell, C. 1988. "Corporate Consumers Affairs Departments: Retrospect and Prospect." Pp. 595–620 in *Frontiers of Research on Consumers Interest*, edited by E. Scott Maynes. Columbia, Mo.: American Council on Consumer Interest.
- Fox, William. 1984. *The Mirror Makers*. New York: William Morrow.
- Frenzen, J., Paul Hirsch, and Phillip Zerrillo. 1994. "Consumption, Preferences, and Changing Lifestyle." Pp. 403–25 in *Handbook of Economic Sociology*, edited by Neil Smelser and Richard Swedberg. New York: Russell Sage.
- Friedland, Roger, and Robert R. Alford. 1991. "Bringing Society Back In: Symbols, Practices, and Institutional Contradictions." Pp. 232–66 in *The New Institutionalism in Organizational Analysis*, edited by Walter W. Powell and Paul J. DiMaggio. Chicago: University of Chicago Press.
- Friedson, Elliot. 1986. *Professional Powers: A Study of the Institutionalization of Formal Knowledge*. Chicago: University of Chicago Press.
- Fuld, Leonhard Felix. 1916. *Service Instructions of American Corporations*. Washington, D.C.: U.S. Department of the Interior, Bureau of Education.
- Granovetter, Mark. 1994. "Business Groups." Pp. 454–76 in *Handbook of Economic Sociology*, edited by N. J. Smelser and R. Swedberg. New York: Russell Sage.
- Haines, Herbert H. 1984. "Black Radicalization and the Funding of Civil Rights: 1957–1970." *Social Problems* 32:31–43.
- Hannan, Michael T., and John Freeman. 1989. *Organizational Ecology*. Cambridge, Mass.: Belknap Press.
- Hechter, Michael T. 1990. "The Emergence of Cooperative Social Institutions." Pp.

- 13–34 in *Social Institutions: Their Emergence, Maintenance, and Effects*, edited by Michael Hechter, Karl-Dieter Opp, and Reinhard Wippler. New York: Aldine de Gruyter.
- Heermance, Edgar L. 1924. *Codes of Ethics: A Handbook*. Burlington, Vt.: Free Press.
- Hermann, Robert O. 1970. "Consumerism. Its Goals, Organizations and Future." *Journal of Marketing* 34 (October): 55–60.
- Hirsch, Paul M. 1986. "From Ambushes to Golden Parachutes: Corporate Takeovers as an Instance of Cultural Framing and Institutional Integration." *American Journal of Sociology* 91:801–37.
- Kallett, Arthur, and Frederick Schlink. 1933. *100,000,000 Guinea Pigs: Dangers in Foods, Drugs, and Cosmetics*. New York: Grosset & Dunlap.
- Katz, Norman. 1977. "Consumers Union: The Movement and the Magazine." Ph.D. dissertation. Rutgers University.
- Kelley, Florence. 1899. "Aims and Principles of the Consumers' League." *American Journal of Sociology* 5 (3): 289–304.
- Kimberly, J. R. 1975. "Environmental Constraints and Organizational Structure: A Comparative Analysis of Rehabilitation Organizations." *Administrative Science Quarterly* 20:1–19.
- Kitson, Harry. 1922. "The Growth of the 'Service Idea' in Selling." *Journal of Political Economy* 30 (June): 417–19.
- Klandermans, B. 1992. "The Social Construction of Protest and Multi-Organizational Fields." Pp. 77–103 in *Frontiers in Social Movement Theory*, edited by Aldon Morris and Carol Mueller. New Haven, Conn.: Yale University Press.
- Lamb, Ruth deForest. 1936. *American Chamber of Horrors: The Truth about Food and Drugs*. New York: Farrar & Rinehart.
- Letts, Arthur. 1914. "Progressive Methods in Dry Goods Retailing." *Dry Goods Economist* 68 (April): 7.
- Luhmann, Niklas. 1979. *Trust and Power: Two Works*. Chichester, N.Y.: Wiley.
- Lukes, S. 1974. *Power A Radical View*. London: Macmillan.
- Lumsden, C., and Jitendra Singh. 1990. "The Dynamics of Organizational Speciation." Pp. 145–63 in *Organisational Evolution: New Directions*, edited by Jitendra V. Singh. San Francisco: Sage.
- Madden, M. Stuart. 1988. *Products Liability*. St. Paul, Minn.: West.
- March, James, and Johan Olsen. 1989. *Rediscovering Institutions*. New York: Free Press.
- McAdam, D. 1994. "Culture and Social Movements." Pp. 36–57 in *New Social Movements*, edited by Enrique Larana, Hank Johnston, and Joseph R. Gusfield. Philadelphia: Temple University Press.
- McAdam, D., J. D. McCarthy, and M. N. Zald. 1988. "Social Movements." Pp. 695–737 in *Handbook of Sociology*, edited by Neil J. Smelser. Beverly Hills, Calif.: Sage.
- McKelvey, B. 1982. *Organisational Systematics: Taxonomy, Evolution, Classification*. Berkeley and Los Angeles: University of California Press.
- Meyer, D. S., and D. R. Imig. 1993. "Political Opportunity and the Rise and Decline of Interest Group Sectors." *Social Science Journal* 30:253–70.
- Meyer, John W., John Boli, and George M. Thomas. 1987. "Ontology and Rationalization in the Western Cultural Account." Pp. 12–37 in *Institutional Structure: Constituting State, Society, and the Individual*, edited by George M. Thomas, John W. Meyer, Francisco Ramirez, and John Boli. Newbury Park, Calif.: Sage.
- Meyer, John, and Brian Rowan. 1991. "Institutional Organizations: Formal Structures as Myth and Ceremony." *American Journal of Sociology* 83:340–63.
- Nader, Ralph. 1965. *Unsafe at Any Speed: The Designed-in Dangers of the American Automobile*. New York: Grossman.
- Nelson, Richard R., and Sidney G. Winter. 1982. *An Evolutionary Theory of Economic Change*. Cambridge, Mass.: Belknap Press.

- Oliver, Christine 1991 "Strategic Responses to Institutional Processes." *Academy of Management Review* 16:145-79.
- . 1992. "The Antecedents of Deinstitutionalization." *Organization Studies* 13: 563-88.
- Olney, M. 1991. *Buy Now, Pay Later*. Chapel Hill: University of North Carolina Press.
- Oster, Sharon. 1994. *Modern Competitive Analysis*. New York: Oxford University Press.
- Padgett, John, and Christopher K. Ansell. 1993. "Robust Action and the Rise of the Medici, 1400-1434." *American Journal of Sociology* 98:1259-1319.
- Pfeffer, J. M. 1992. *Managing with Power*. Cambridge, Mass.: Harvard University Press.
- Phillips, Mary C. 1934. *Skin Deep*. New York: Vanguard.
- Pollay, R. 1985. "The Subsidizing Sizzle: A Descriptive History of Print Advertising, 1900-1980." *Journal of Marketing* 49 (3): 24-39.
- Powell, Walter W. 1991. "Expanding the Scope of Organizational Analysis." Pp 183-203 in *The New Institutionalism in Organizational Analysis*, edited by Walter W. Powell and Paul J. DiMaggio. Chicago: University of Chicago Press.
- Priest, G. 1984. "The Invention of Enterprise Liability: A Critical History of the Intellectual Foundations of Modern Tort Law." Working paper, no. 27. Yale Law School, Yale University.
- Ranson, S., C. L. Hinings, and Royston Greenwood. 1980. "The Structuring of Organizational Structure." *Administrative Science Quarterly* 25:1-14.
- Romanelli, E. 1989. "Organizational Birth and Population Variety: A Community Perspective on Origins." Pp 211-246 in *Research in Organizational Behavior*, vol. 11. Edited by Barry Staw and Larry L. Cummings. Greenwich, Conn.: JAI Press.
- Samson, Peter. 1980. "The Emergence of a Consumer Interest in America." Ph D dissertation. University of Chicago.
- Sargent, Hugh. 1958. "The Influence of Consumer Product Testing and Reporting Services on Consumer Behavior." Ph.D. dissertation. University of Illinois.
- Schlink, F. 1927. "Tests and Specifications for the Household." *Journal of Home Economics* 19:18-20.
- Schudson, Michael. 1978. *Discovering the News*. New York: Basic.
- Schumpeter, J. A. 1950. *Capitalism Socialism and Democracy*. New York: Harper.
- Scott, W. R. 1995. *Institutions and Organizations*. San Francisco: Sage.
- Shapiro, Susan P. 1987. "The Social Control of Impersonal Trust." *American Journal of Sociology* 93:623-58.
- Shepsle, K. 1990. *Perspectives on Positive Political Economy*. Cambridge: Cambridge University Press.
- Silber, Norman Isaac. 1983. *Test and Protest: The Influence of Consumers Union*. New York: Holmes & Meier.
- Sorenson, Helen. 1941. *The Consumer Movement*. New York: Harper.
- Snow, David A., and Robert D. Benford. 1992. "Master Frames and Cycles of Protest." Pp. 133-55 in *Frontiers in Social Movement Theory*, edited by Aldon Morris and Carol Mueller. New Haven, Conn.: Yale University Press.
- Spaone, C. A. 1985. "The Emergence of Strict Liability: A Historical Perspective and Other Considerations including Senate 100." *Journal of Product Liability* 8:1-40.
- Stinchcombe, Arthur L. 1965 "Social Structure and Organizations." Pp. 142-93 in *Handbook of Organizations*, edited by James G. March. Chicago: Rand McNally.
- . 1968. *Constructing Social Theories*. Chicago: University of Chicago Press.
- Stoddard, C. K. 1928. "Your Standard's Worth." *Printers Ink*, August 16, pp. 117-20.
- Sutton, J. R., and F. Dobbin. 1996. "The Two Faces of Governance: Responses to Legal Uncertainty in U.S. Firms, 1955-1986." *American Sociological Review* 61: 794-811.

- Swidler, Ann. 1986. "Culture in Action. Symbols and Strategies." *American Sociological Review* 51:273-86.
- Tarrow, Sidney. 1989. *Democracy and Disorder: Protest and Politics in Italy: 1965-1975*. Oxford: Oxford University Press.
- Thomas, Robert Paul. 1977. *An Analysis of the Patterns of Growth of the Automobile Industry, 1895-1929*. New York: Arno Press.
- Thorelli, Hans, and Sarah Thorelli. 1974. *Consumer Information Handbook*. New York: Praeger.
- Tushman, Michael L., and Phillip Anderson. 1986. "Technological Discontinuities and Organizational Environments." *Administrative Science Quarterly* 31:439-65.
- U.S. Bureau of the Census. 1975. *Historical Statistics of the United States, Colonial Times to 1970*. Washington, D.C.: U.S. Department of Commerce.
- Vaile, R. 1927 *The Economics of Advertising*. New York: Ronald Press.
- Vandall, F. 1989. *Strict Liability: Legal and Economic Analysis*. New York: Quorum Books.
- Van de Ven, A. H., and R. Garud. 1989. "A Framework for Understanding the Emergence of New Industries." Pp. 195-225 in *Research In Technological Innovation, Management, and Policy*, edited by R. Rosenbloom and R. Burgelman. Greenwich, Conn.: JAI Press.
- Veblen, Thorstein. 1919. *The Place of Science in Modern Civilization and Other Essays*. New York: Huebsch.
- White, Harrison. 1992. *Identity and Control*. Princeton, N.J.: Princeton University Press.
- Wiley, Mary, and Mayer Zald. 1968. "The Growth and Transformation of Educational Accrediting Agencies: An Exploratory Study in the Social Control of Institutions." *Sociology of Education* 41 36-56.
- Williamson, Oliver E. 1985. *The Economic Institutions of Capitalism*. New York: Free Press.
- Zald, Mayer N. 1992. "Looking Backward to Look Forward: Reflections on the Past and Future of the Resource Mobilization Research Program." Pp. 326-48 in *Frontiers in Social Movement Theory*, edited by Aldon Morris and Carol Mueller. New Haven, Conn.: Yale University Press.
- Zald, Mayer, and John D. McCarthy. 1980. "Social Movement Industries: Competition and Cooperation among Movement Organizations." *Research in Social Movements, Conflict, and Change* 3:1-20.
- Zelizer, Vivlianna A. 1994. *The Social Meaning of Money*. New York: Basic.
- Zucker, Lynne G. 1986 "The Production of Trust: Institutional Sources of Economic Structure: 1840-1920." Pp. 53-111 in *Research in Organizational Behavior*, vol. 8. Edited by Barry Staw and Larry L. Cummings. Greenwich, Conn.: JAI Press.
- Zuckerman, E. 1996 "Mediating the Corporate Product: Securities Analysts and the Scope of the Firm." Working paper. University of Chicago, Department of Sociology.

What Is Agency?¹

Mustafa Emirbayer and Ann Mische
New School for Social Research

This article aims (1) to analytically disaggregate agency into its several component elements (though these are interrelated empirically), (2) to demonstrate the ways in which these agentic dimensions interpenetrate with forms of structure, and (3) to point out the implications of such a conception of agency for empirical research. The authors conceptualize agency as a temporally embedded process of social engagement, informed by the past (in its "iterational" or habitual aspect) but also oriented toward the future (as a "projective" capacity to imagine alternative possibilities) and toward the present (as a "practical-evaluative" capacity to contextualize past habits and future projects within the contingencies of the moment).

The concept of agency has become a source of increasing strain and confusion in social thought. Variants of action theory, normative theory, and political-institutional analysis have defended, attacked, buried, and resuscitated the concept in often contradictory and overlapping ways. At the center of the debate, the term *agency* itself has maintained an elusive, albeit resonant, vagueness; it has all too seldom inspired systematic analysis, despite the long list of terms with which it has been associated: selfhood, motivation, will, purposiveness, intentionality, choice, initiative, freedom, and creativity. Moreover, in the struggle to demonstrate the interpenetration of agency and structure, many theorists have failed to dis-

¹ This is a fully coauthored article. Earlier drafts were presented at the Paul F. Lazarsfeld Center for the Social Sciences at Columbia University, the Workshop on Politics, Power, and Protest at New York University, the Colloquium on Culture and Politics at the New School for Social Research, the meeting of the American Sociological Association at Los Angeles, and various seminars at the New School for Social Research and Princeton University. We would like to thank the participants in those forums for their many useful comments. We would also like to thank Jeffrey Alexander, Bernard Barber, Richard Bernstein, Donald Black, Mary Blair-Loy, David Gibson, Chad Goldberg, Jeff Goodwin, Michael Hanagan, Hans Joas, Michèle Lamont, Edward Lehman, Calvin Morrill, Michael Muhlhaus, Shepley Orr, Margarita Palacios, Mimi Sheller, Charles Tilly, Diane Vaughan, Loïc Wacquant, and Harrison White for their many illuminating insights, criticisms, and suggestions. Direct correspondence to Mustafa Emirbayer, Department of Sociology, New School for Social Research, 65 Fifth Avenue, New York, New York 10003

tinguish agency as an analytical category in its own right—with distinctive theoretical dimensions and temporally variable social manifestations. The result has been a flat and impoverished conception that, when it escapes the abstract voluntarism of rational choice theory, tends to remain so tightly bound to structure that one loses sight of the different ways in which agency actually shapes social action.

We argue that each of the most significant recent attempts to theorize agency has neglected crucial aspects of the problem. In distinguishing (and showing the interplay) between different dimensions of agency, we seek to go beyond these various one-sided points of view. "Theorists of practice" such as Pierre Bourdieu and Anthony Giddens, for example, have given selective attention to the role of habitus and routinized practices; their perspective (perhaps the dominant one in contemporary American sociology) sees human agency as habitual, repetitive, and taken for granted—a view shared by ethnomethodologists, new institutionalists in organizational theory, and many others. Alternative approaches have similarly relied upon one-sided conceptions of agency; for example, traditions as different from one another as rational choice theory and phenomenology have stressed goal seeking and purposivity, while theories of publicity and communication, as well as certain feminist theories, have overemphasized deliberation and judgment. While routine, purpose, and judgment all constitute important dimensions of agency, none by itself captures its full complexity. Moreover, when one or another is conflated with agency itself, we lose a sense of the dynamic *interplay* among these dimensions and of how this interplay varies within different structural contexts of action.

Our immediate aims in this article, then, are threefold: (1) to analytically disaggregate agency into its several component elements (even though these are clearly interrelated empirically), (2) to demonstrate the different ways in which the dimensions of agency interpenetrate with diverse forms of structure, and (3) to point out the implications of such a differentiated conception of agency for empirical research.

Theoretically, our central contribution is to begin to reconceptualize human agency as a temporally embedded process of social engagement, informed by the past (in its habitual aspect), but also oriented toward the future (as a capacity to imagine alternative possibilities) and toward the present (as a capacity to contextualize past habits and future projects within the contingencies of the moment). The agentic dimension of social action can only be captured in its full complexity, we argue, if it is analytically situated within the flow of time. More radically, we also argue that the structural contexts of action are themselves temporal as well as relational fields—multiple, overlapping *ways of ordering time* toward

which social actors can assume different simultaneous agentic orientations. Since social actors are embedded within many such temporalities at once, they can be said to be oriented toward the past, the future, and the present at any given moment, although they may be primarily oriented toward one or another of these within any one emergent situation. As actors move within and among these different unfolding contexts, they switch between (or "recompose") their temporal orientations—as constructed within and by means of those contexts—and thus are capable of changing their relationship to structure. We claim that, in examining changes in agentic orientation, we can gain crucial analytical leverage for charting varying degrees of maneuverability, inventiveness, and reflective choice shown by social actors in relation to the constraining and enabling contexts of action.

Most broadly, our guiding concerns in this article are moral and practical in nature. We contend that reconceptualizing agency as an internally complex temporal dynamic makes possible a new perspective upon the age-old problem of free will and determinism. How are social actors, we ask, capable (at least in principle) of critically evaluating and reconstructing the conditions of their own lives? If *structural contexts* are analytically separable from (and stand over against) capacities for *human agency*, how is it possible for actors ever to mediate or to transform their own relationships to these contexts? Without disaggregating the concept of agency into its most important analytical dimensions, we cannot ever hope to find satisfactory answers to these questions. The key to grasping the dynamic possibilities of human agency is to view it as composed of variable and changing orientations within the flow of time. Only then will it be clear how the structural environments of action are both dynamically sustained by and also altered through human agency—by actors capable of formulating projects for the future and realizing them, even if only in small part, and with unforeseen outcomes, in the present.

THEORIZING AGENCY

Many of the tensions in present-day conceptions of human agency can be traced back to the Enlightenment debate over whether instrumental rationality or moral and norm-based action is the truest expression of human freedom. Teleological and instrumentalist conceptions of action fueled the philosophical individualism of the early Enlightenment, which, while still grounded in the religious morality of the times, allowed for the subsequent invention of the individual as a "free agent" able to make rational choices for (him)self and society (Lukes 1973). With John Locke's (1978) rejection of the binding power of tradition, his location of beliefs in individual experience, and his grounding of society in the social contract

between individuals, a new conception of agency emerged that affirmed the capacity of human beings to shape the circumstances in which they live. This faith subsequently sustained a long line of social thinkers, including Adam Smith, Jeremy Bentham, and John Stuart Mill, and embedded agency in an individualist and calculative conception of action that still underlies many Western accounts of freedom and progress.

In response to this association of freedom with rational self-interest, other Enlightenment thinkers, most notably Jean-Jacques Rousseau, anticipated the later Romantics by exploring instead such alternative conceptions of freedom as the ascendancy of conscience and moral will, of a self-legislating morality. Their perspective underscored the importance of the transcendental imagination as well as that of instrumental reason. These two points of view both found their way into Immanuel Kant's (1765, 1784, 1797) critical philosophy, which saw freedom as normatively grounded individual will, governed by the categorical imperative rather than by material necessity (or interest). Kant bifurcated all of reality into two opposing orders: the conditional and the normative, necessity and freedom—the latter conceived of as the pure unconditioned activity of autonomous moral beings. His rendering of the ancient question of free will versus necessity became in classical sociological theory the point of departure for a concern with nonrational norm-oriented action—in contradistinction to the rational instrumental action emphasized by economic analysts of society (Habermas 1984–89; Münch 1981, 1994). In Hans Joas's (1993, p. 247) words, "As a safeguard against the utilitarian dangers of the theory of rational action, the founding theorists of sociology [had] recourse to Kant and his notion of free, moral action." In this line, the early action theory of Talcott Parsons can be read as a Kant-inspired attempt to synthesize the rational-utilitarian and nonrational-normative dimensions of action. In *The Structure of Social Action*, for example, Parsons (1968, p. 732) argued that "conditions may be conceived at one pole, ends and normative rules at another, means and effort as the connecting link between them." Agency, for Parsons, was captured in the notion of *effort*, as the force that achieves, in Kantian terminology, the interpenetration of means-ends rationality and categorical obligation.

Parsons's early attention to the temporal dimension of action (subsequently discarded in his later structural-functionalist work) also remained caught within Kantian dualisms. He noted that all social action, whether instrumental or normative, is teleological in structure: "An act is always a process in time. . . . The concept end always implies a future reference, to a state which is . . . not yet in existence, and which would not come into existence if something were not done about it by the actor" (Parsons 1968, p. 45). In none of his writings, on the other hand, did Parsons elaborate a fully temporal theory of agency (or, indeed, of structure): agency

remained "outside" of time (as in Kant's own conception of the "unconditioned"), while structure remained a spatial category rather than (also) a temporal construction. Moreover, in none of his writings did Parsons devote much systematic attention to disaggregating the crucial concept of effort itself—to opening up the "black box" of human agency.

Agency in Social Theory

In explicit dialogue with Parsonian (and Kantian) theories of agency, both James Coleman and Jeffrey Alexander have recently presented attempts to join instrumental and normative approaches, although with strikingly different results. Responding to the disappearance of agency in later versions of structural-functionalism, rational choice advocates have followed George Homans's (1964) call to "bring men back in" and to return to an action theory firmly grounded in the purposive, instrumental, and calculating orientations of individuals. In his major synthetic work, *Foundations of Social Theory*, Coleman (1990) tries to overcome the Kantian division between interests and norms by arguing that rational choice assumptions can provide the underpinnings for a normative theory based upon power-weighted social influence. Coleman counters the decontextualized individualism of many rational actor perspectives by linking purposive activity at the micro level to systemic interdependencies at the macro level, thereby showing that action is always a complex social and interactive phenomenon. However, he fails to address the problem at the heart of rational choice explanations: the (clearly acknowledged) decision to bracket the question of how temporally embedded actors actually reach decisions that can retrospectively be interpreted as rational. By assuming that "actions are 'caused' by their (anticipated) consequences" Coleman (1986, p. 1312) attributes the impulse to action to a means-ends rationality abstracted from the human experience of time. While this bracketing of subjective temporality does in fact lead to the prediction of an impressive range of social phenomena resulting from individual choices, it does not allow us to understand the interpretive processes whereby choices are imagined, evaluated, and contingently reconstructed by actors in ongoing dialogue with unfolding situations. The post hoc causal attribution implicit in rational choice conceptions of agency leaves Parsons's black box untouched.¹

¹ We acknowledge that many rational choice theorists have made great strides in accounting for the contingencies and uncertainties involved in choice making (March and Simon 1958; March and Olsen 1976; March 1978), as well as in attempting to explore the role of values, norms, and other cultural elements (Elster 1989; Hechter 1992, 1994; see also the essays in Cook and Levi [1990]). However, we maintain that even these more sophisticated versions of rational actor models are still grounded in presuppositions that prevent them from adequately theorizing the interpretive inter-

A more promising initiative in the analytic exploration of agency can be seen in the recent work of Jeffrey Alexander (1988, esp. pp. 301–33; 1992). Although a neo-Parsonian himself in many respects, and thus influenced in the deep structure of his thought by Kantian categories (he continues to take as his frame of reference the dichotomy between the conditional and the normative), Alexander advances considerably beyond both Kant and Parsons in thematizing the ways in which human agency engages with its structural contexts. He is the first major theorist to systematically disaggregate the concept of agency itself, probing into its inner structure and delineating categories of agentic processes. In *Action and Its Environments*, Alexander (1988) proposes that action be conceived of in terms of two basic dimensions, which he calls *interpretation* (further subdivided into *typification* and *invention*) and *strategization*. He intends by these analytical categories to synthesize, as did Parsons before him, the normative and utilitarian perspectives by presenting them as complementary but analytically distinguishable dimensions of human action. But Alexander's multidimensional theory also goes much further than Parsonian theory in providing insight into precisely that element bracketed by Coleman, that is, the interpretive processes of contextually embedded actors. In what follows, we build upon Alexander's highly useful categorization, which opens up theoretical space for analyzing the inventive and critical aspects of agency. We contend, however, that because his analysis remains subsumed under a broader category of normativity, he has little to say about invention's constitutive features and, specifically, its pragmatic and experimental dimensions. Even more important, Alexander neglects to situate his analysis of agency within a specifically temporal framework. We argue, by contrast, that agentic processes can only be understood if they are linked intrinsically to the changing temporal orientations of situated actors.

To place agency within such a temporal framework, and to move effectively beyond the division between instrumental and normative action, we must turn to the philosophical school that most consistently challenges such dualisms, notably American pragmatism (with its close ties to Continental phenomenology). In response to the utilitarian model of rational action, pragmatist thinkers such as John Dewey and George Herbert Mead, as well as social phenomenologists such as Alfred Schutz, insist that action not be perceived as the pursuit of preestablished ends, abstracted from concrete situations, but rather that ends and means develop coterminously within contexts that are themselves ever changing and thus always subject to reevaluation and reconstruction on the part of the re-

subjective construction of choices from the temporal vantage points of contextually embedded actors.

flective intelligence. Moreover, pragmatists reject the Kantian response to utilitarianism by condemning the false distinction between material interests and transcendental values, since all human objects and purposes are necessarily constructed out of social meanings and values. These basic premises allow the pragmatist thinkers to sidestep many of the conundrums that dominate sociological thought and to lay the foundations for a theory of action that analyzes the "conditions of possibility" (Joas 1993, p. 250) for the evaluative, experimental, and constructive dimensions of perception and action, within the contexts of social experience.

While we draw upon a variety of pragmatist and phenomenological thinkers in the sections to come, it is the work of George Herbert Mead that offers us the most compelling tools for overcoming the inadequate conceptions of agency in both rational choice and norm-oriented approaches. Although Mead is best known for his contributions to social psychology and symbolic interactionism, we focus here upon his seminal (but little discussed) theorization of temporality in *The Philosophy of the Present* (1932).³ Two insights in this work are critical for our efforts: the concept of time as constituted through *emergent events*, which require a continual refocusing of past and future, and the concept of human consciousness as constituted through *sociality*, the capacity to be both *temporally* and *relationally* in a variety of systems at once. Building upon the work of Henri Bergson (1989), Mead rejects the Newtonian conception of time as a succession of isolated instants, characterizing time instead as a multilevel flow of nested events, radically grounded in (but not bounded by) present experience. "Reality exists in a present" (Mead 1932, p. 1), although the immediacy of present situations is extended by our ability to imaginatively construct a sense of past and future. But Mead also moves beyond the individualist and subjectivist presuppositions of Bergson's theory, which conceptualizes time as an introspective *durée*, a merely psychological rather than intrinsically social phenomenon. By contrast, Mead insists that the human experience of temporality is based in the *social* character of emergence, that is, in the passage from the old to the new, and in the interrelated changes occurring throughout the various situational contexts within which human beings are embedded. As actors respond to changing environments, they must continually reconstruct

³ We are not concerned here with Mead's engagement in this work with functionalist evolutionary theory nor with his debate with metaphysical theorists of temporality. Although Mead develops his theories through a comparison with more general physical and biological (i.e., nonhuman) processes and has been criticized for veering away from action theory toward metaphysics (Joas 1985), he also provides the philosophical core of a temporal and relational understanding of the intersubjective development of agentic capacities, which is of critical importance for a theory of action. For a related discussion, see also Mead's (1938) work, *The Philosophy of the Act*.

their view of the past in an attempt to understand the causal conditioning of the emergent present, while using this understanding to control and shape their responses in the arising future. This process forms the core of what Mead (1932, p. 76) calls "the deliberative attitude," the capacity to "get hold of the conditions of future conduct as these are found in the organized responses we have formed, and so construct our pasts in anticipation of that future."

Mead points this insight in the direction of action theory by describing how what he calls sociality—that is, the situatedness of actors in multiple temporally evolving relational contexts—contributes to the development of reflective consciousness. Mead outlines three levels of consciousness, distinguished in terms of the increasing capacity of actors to actively constitute their environments through selective control over their own responses: (1) the level of "contact experience," characterized by immediacy of response to sense and feeling, (2) that of "distance experience," characterized by the capacity to use ideation and imagery in remembrance and anticipation, and finally, (3) the culmination of sociality in communicative interaction, in which social meanings and values develop out of the capacity to take on the perspectives of (concrete and generalized) others. What drives the development of consciousness from one level to the next is the "awakening of delayed and conflicting responses" (Mead 1932, p. 71) to problematic situations in one's various environments, increasing the field of choice while extending the temporal perspective of action. At every step, actors are conceived of not as atomized individuals, but rather as active respondents within nested and overlapping systems (which we prefer to call *temporal-relational contexts*); the construction of temporal perspectives is fundamentally an *intersubjective* process, constituted by the ability to hold simultaneously to one's own and to another's viewpoint. Actors develop their deliberative capacities as they confront emergent situations that impact upon each other and pose increasingly complex problems, which must be taken up as challenges by the responsive (and communicative) intelligence.

Unlike Mead, we are not primarily interested in the *evolution* of reflective consciousness but rather in the insight that Mead's analysis affords into the internal structuring of agentic capacities and their different constitutive relationships to action. We agree with Hans Joas in his recent book, *The Creativity of Action* (1996; see also Joas, n.d.), that pragmatist thinkers provide the first steps toward developing an adequate conception of the constitutive creativity of action, conceived of as "the permanent reorganization and reconstitution of habits and institutions" (Joas, n.d., p. 24). Such a conception, Joas argues, fundamentally challenges the teleological means-ends model present in both rational choice and neo-Parsonian approaches, replacing it with an account of the situational and

corporeal embeddedness of action.⁴ Joas's major contribution is to wrest the theory of action from both its rationalist and norm-centered pre-suppositions, insisting that a conception of the situationally embedded creativity of action is essential not only for studies of microinteraction, but also for macrosociological analysis (and particularly for understanding the possibilities of what Dewey calls *creative democracy*). Yet he brackets the major question that we examine here, that of "large differences in the various acts and actors in regards to creativity" (Joas 1996, p. 197). We maintain that this is not merely an empirical but also an analytical question: by differentiating between the different dimensions of agency, we can help to account for variability and change in actors' capacities for imaginative and critical intervention in the diverse contexts within which they act.

The Chordal Triad of Agency

What, then, is human agency? We define it as *the temporally constructed engagement by actors of different structural environments—the temporal-relational contexts of action—which, through the interplay of habit, imagination, and judgment, both reproduces and transforms those structures in interactive response to the problems posed by changing historical situations*.⁵ This definition encompasses what we shall analytically distinguish below as the different constitutive elements of human agency: iteration, projectivity, and practical evaluation. In broad terms, these correspond

⁴ For Joas (1996, p. 160), action is not simply *contingent* upon the situation, but more essentially, "the situation is *constitutive* of action" (original emphasis), providing not merely "means" and "conditions" for preestablished ends but also the structured habitual patterns of response that become the basis for the reflective and creative engagement of actors with their changing environments.

⁵ While our principal focus in this article remains the different analytical dimensions of agency rather than action's structural contexts, we follow earlier work (Emirbayer and Goodwin 1996)—along with Sorokin (1947), Parsons and Shils (1951), and, especially, Alexander (1988b)—in our disaggregation of the latter. As we conceive of it, the *cultural context* encompasses those symbolic patterns, structures, and formations (e.g., cultural discourses, narratives, and idioms) that constrain and enable action by structuring actors' normative commitments and their understandings of their world and their possibilities within it. The *social-structural context* encompasses those network patterns of social ties (see Emirbayer and Goodwin 1994) that comprise interpersonal, interorganizational, or transnational settings of action. Finally, the *social-psychological context* encompasses those psychical structures that constrain and enable action by channelling actors' flows and investments of emotional energy, including long-lasting durable structures of attachment and emotional solidarity. These interpenetrating (but analytically autonomous) categories crosscut the key institutional sectors of modern social life: the administrative-bureaucratic state, the capitalist economy, and civil society (Emirbayer and Sheller 1996).

to the different temporal orientations of agency, allowing us to examine forms of action that are more oriented (respectively) toward the past, the future, and the present. Such a categorization gives analytical expression to Mead's conception of the positioning of human actors within temporal passage, involving the continual reconstruction of their orientations toward past and future in response to emergent events. In addition, it incorporates Mead's insight that it is the capacity for imaginative distancing, as well as for communicative evaluation, in relation to habitual patterns of social engagement that drives the development of the reflective intelligence, that is, the capacity of actors to critically shape their own responsiveness to problematic situations.

The iterational element.—The first of these dimensions, which we term the iterational element, has received perhaps the most systematic attention in philosophy and sociological theory, most recently from that tradition of thought that Ortner (1984) describes as *theories of practice* (see also Turner 1994). It refers to *the selective reactivation by actors of past patterns of thought and action, as routinely incorporated in practical activity, thereby giving stability and order to social universes and helping to sustain identities, interactions, and institutions over time.*

The projective element.—The second dimension of agency, the projective element, has been largely neglected in recent sociological theory, although it does receive attention in the writings of Alfred Schutz and his followers, and, indirectly, of rational choice theorists. Outside of sociology, concern with projectivity can be found in phenomenological and existential philosophy, psychoanalysis, narrative psychology, and dramaturgic anthropology. Projectivity encompasses *the imaginative generation by actors of possible future trajectories of action, in which received structures of thought and action may be creatively reconfigured in relation to actors' hopes, fears, and desires for the future.*

The practical-evaluative element.—Finally, the practical-evaluative element of agency has been left strikingly undertheorized by sociological thinkers, although intimations of it can be found in a long tradition of moral philosophy extending from Aristotelian ethics to more recent theories of critical deliberation, as well as certain feminist analyses. It entails *the capacity of actors to make practical and normative judgments among alternative possible trajectories of action, in response to the emerging demands, dilemmas, and ambiguities of presently evolving situations.*

We should stress from the outset that these are analytical distinctions; all three of these constitutive dimensions of human agency are to be found, in varying degrees, within any concrete empirical instance of action. In

this sense, it is possible to speak of a *chordal triad* of agency within which all three dimensions resonate as separate but not always harmonious tones.⁶ On the other hand, we also claim that, in any given case, one or another of these three aspects might well predominate. It is possible to speak of action that is more (or less) engaged with the past, more (or less) directed toward the future, and more (or less) responsive to the present. In each of the three major sections below, we isolate these various analytical dimensions and examine the internal structure of each. Although it will never be possible to carry out our analytical dissections with surgical precision, we aim to show what agentic processes would entail were one or another of these tones in the chordal triad to be sounded most forcefully.⁷

Moreover, we also argue that each of the three analytical dimensions can be said to possess its own internal chordal structure. The three dimensions of agency that we describe do not correspond in any simple, exclusive way to past, present, and future as successive stages of action. Rather, empirical social action is constructed through ongoing temporal passage and thus through what Mead calls emergent events, rather than through a sequentiality of discrete acts or stages of one act. Each of our dimensions of agency has itself a simultaneous internal orientation toward past, future, and present, for all forms of agency are temporally embedded in the flow of time. We do claim, however, that for each analytical aspect of agency one temporal orientation is the dominant tone, shaping the way in which actors relate to the other two dimensions of time. Disaggregating the dimensions of agency (and exploring which orientations are dominant within a given situation) allows us to suggest that each primary orientation in the chordal triad encompasses as subtones the other two as well, while also showing how this "chordal composition" can change as actors respond to the diverse and shifting environments around them.⁸

Several further points of clarification are in order here. First, we must reaffirm that agency as we have sketched it above is a historically variable

⁶ This usage is analogous to Patterson's (1991) discussion of the chordal triad of freedom.

⁷ We bracket for now the added complication that actors are always embedded within many different temporal-relational contexts at once and thus may exhibit a projective orientation within one context, e.g., even as they exhibit an iterational orientation within another. We return to this issue in the final section of the article.

⁸ Lest we fall into the analytical nightmare of "subsubtones" within "subtones," we wish to stress that the notion of an internal chordal structure is a heuristic device that allows us to analyze variation and change in the composition of agentic orientations; clearly, actors do not dissect experience in such a manner while themselves in the flow of temporal passage. We should also note that what we call *chordal structures* are not necessarily harmonious; the subtones may be dissonant with one another, creating internal tensions that may spur the recomposition of temporal orientations.

phenomenon, embedded in changing theoretical and practical conceptions of time and action. Ours is not a universalistic perspective that assumes that all times, places, and persons are equally iterational, projective, or practical-evaluative. Rather, it is precisely the historical, cultural, and personal variability of agentic orientations that make this framework so compelling. The ways in which people understand their own relationship to the past, future, and present *make a difference* to their actions; changing conceptions of agentic possibility in relation to structural contexts profoundly influence how actors in different periods and places see their worlds as more or less responsive to human imagination, purpose, and effort.

Second, we follow Mead in arguing that changes in temporal orientation may also involve varying degrees of inventiveness and reflectivity in relation to action and its temporal-relational contexts, although not necessarily, as we shall show later, in simple or straightforward ways. (Such a conception signals our deliberate commitment to a humanistic, normative, and critical perspective upon social life.) While we claim that even habitual action is agentic, since it involves attention and effort, such activity is largely unreflective and taken for granted; as actors encounter problematic situations requiring the exercise of imagination and judgment, they gain a reflective distance from received patterns that may (in some contexts) may allow for greater imagination, choice, and conscious purpose. A disaggregated conception of agency thus allows us to locate more precisely the interplay between the reproductive and transformative dimensions of social action (Hays 1994) and to explain how reflectivity can change in either direction, through the increasing routinization or problematization of experience.

Third, we wish to stress that our conception of agency is intrinsically social and relational (Emirbayer 1997) since it centers around the engagement (and disengagement) by actors of the different contextual environments that constitute their own structured yet flexible social universes. For this reason, and also because of our deep resonance with both classical and contemporary pragmatism, one might characterize our approach as *relational pragmatics*. Viewed internally, agency entails different ways of experiencing the world, although even here, just as consciousness is always consciousness of something (James 1976; Husserl 1960), so too is agency always agency *toward* something, by means of which actors enter into relationship with surrounding persons, places, meanings, and events. Viewed externally, agency entails actual interactions with its contexts, in something like an ongoing conversation; in this sense, it is "filled with dialogic overtones," as a sort of "link in the chain of speech communication" (Bakhtin 1986, pp. 92, 91). Following Mead and Joas, we highlight the importance of intersubjectivity, social interaction, and communication

as critical components of agentic processes: agency is always a dialogical process by and through which actors immersed in temporal passage engage with others within collectively organized contexts of action.

Finally, we ground this capacity for human agency in the structures and processes of the human self, conceived of as an internal conversation possessing analytic autonomy vis-à-vis transpersonal interactions. We conceptualize the self not as a metaphysical substance or entity, such as the "soul" or "will" (see White 1995), but rather as a dialogical structure, itself thoroughly relational. Our perspective, in other words, is relational all the way down.⁹ We cannot begin to explore here the ontology of the self or the full implications for agency of such categories as "desire" (although see Lacan 1977). Nor can we present here a systematic analysis of the components or structures of this self, or elaborate a new philosophical psychology, although we can suggest, following Norbert Wiley (1994, p. 210) in *The Semiotic Self*, that "the interpretive process [taking place within it] is, within limits, open and free," and that this "in turn allows humans to create as well as to pursue goals."¹⁰ We maintain that while transpersonal contexts do both constrain and enable the dialogical process, such contexts cannot themselves serve as the point of origin of agentic possibilities, which must reside one level down (so to speak), at the level of self-dynamics.

In the following discussion, then, we take up in turn three constituent elements of human agency: the iterational, projective, and practical-evaluative tones of the chordal triad. Within each of the sections to come, we first review briefly the relevant history of concepts, then analyze from within the dimension of agency at hand, then finally explore the implications of each aspect for concrete empirical research. In the final major section of the article, we step back to discuss the different ways in which these three dimensions of human agency interpenetrate with different structuring contexts of action.

⁹ Such a position does present us with a certain difficulty: namely, that corporate actors such as firms, states, or other organizational entities cannot easily be accommodated within the terms of such a framework unless they are themselves given theoretical status equivalent to that of natural persons or selves (for examples of this mode of reasoning, see Coleman [1990], Luhmann [1990], and White [1992]). While not averse to such a move in principle, we do not pursue all of its many implications in the pages to come, or grapple systematically with the special challenges in translation that it would necessarily entail.

¹⁰ It is worth noting that Wiley's perspective is itself self-consciously grounded in the pragmatist tradition (see also Wiley 1994, pp. 10, 29, 47); for a similar perspective, see Taylor (1991), Colapietro (1990), and Gergen (1994). More work needs to be done, of course, in theorizing the systematic blockages to such "open and free" intrapsychic communication or dialogue.

THE ITERATIONAL DIMENSION OF AGENCY

If we think of agency as a chordal triad composed of three analytically distinct elements (oriented variously toward the past, future, and present), then what we call the *iterational* dimension appears as that chordal variation in which the *past* is the most resonant tone. Although, as Mead (1932, p. 17) reminds us, all experience takes place in the present, this present is permeated by the conditioning quality of the past: "Its presence is exhibited in memory, and in the historical apparatus that extends memory." Past experiences condition present actions "when they have taken on the organized structure of tendencies" (Mead 1932, p. 18). In this section, we examine how the past, through habit and repetition, becomes a stabilizing influence that shapes the flow of effort and allows us to sustain identities, meanings, and interactions over time. The primary locus of agency for the iterational dimension, we argue, lies in the *schematization* of social experience. It is manifested in actors' abilities to recall, to select, and to appropriately apply the more or less tacit and taken-for-granted schemas of action that they have developed through past interactions. Schemas are corporeal and affective as well as cognitive patterns; they consist in the interpenetration of mental categories, embodied practices, and social organization. Moreover, they constitute temporal as well as relational patterns, recursively implemented in social life (Giddens 1984). The agentic dimension lies in *how actors selectively recognize, locate, and implement* such schemas in their ongoing and situated transactions. While this may take place at a low level of conscious reflection, it still requires attention and engagement on the part of actors in order to narrow the possibilities for action within particular temporal-relational contexts.

The concept of iteration is crucial for our conception of agency since we maintain that both the projective and practical-evaluative dimensions are deeply grounded in habitual, unreflected, and mostly unproblematic patterns of action by means of which we orient our efforts in the greater part of our daily lives. We have settled upon the unfamiliar term *iteration* to describe such activity precisely because the dimension of agency to which it refers is the most difficult to conceive of in properly agentic terms. The subset of words with which it is colloquially associated—routines, dispositions, preconceptions, competences, schemas, patterns, typifications, and traditions—seem more to imply structure than what we commonly think of as agency. This problem is also reflected in most attempts to theorize the habitual dimension of action since they focus upon recurring patterns of action themselves and thus upon structures, rather than upon the precise ways in which social actors relationally engage with those preexisting patterns or schemas.

Iteration: The History of a Concept

In much of social and psychological theory, habit has unfortunately been seen as little more than a matter of stimulus and response, an orientation that shifts attention away from human agency and toward the structural contexts that shape action. Indeed, as Charles Camic (1986, p. 1046) points out, a prevailing tendency in much of social science since the early 20th century has been to regard habit as "behavior that consists in a fixed, mechanical reaction to particular stimuli and [that] is, as such, devoid of meaning from the actor's point of view." The outcome has effectively been to remove habit from the domain of social action.¹¹ In what follows, by contrast, our key concern is to locate the agentic dimension in even the most routinized, prestructured forms of social action. Even relatively unreflective action has its own moment of effort; the typification and routinization of experience are active processes entailing selective reactivation of received structures within expected situations, dynamic transactions between actor and situation. We follow a current of thought (also documented by Camic) that never did succumb to the aforementioned tendency to conceive of habit as a "fixed, mechanical reaction to stimuli" (Camic 1986, p. 1046). According to this perspective, habit entails much more than biophysiological (or institutional) processes; it includes as well the element of agency—no less than do the more reflective and deliberative modes of action.

Classical and medieval philosophy.—Some of the earliest systematic thinking on the iterational aspect of human agency can be found in Aristotle (1985, p. 44), who uses the term *hexis* to refer to any settled disposition or state leading to action. Aristotle distinguishes the *hexis*—sometimes also translated as *habit*—from mechanical behavior as such, since it also reflects a person's desires and decisions. In the *Nicomachean Ethics*, Aristotle (1985) further depicts habits as the basis for "virtues" or "excellences" of character, which entail a settled disposition toward appropriate action in accordance with wisdom. Habits could not form the basis for virtue if they were merely automatic activity. St. Thomas Aquinas, too, defines iterational activity (in his terminology, the *habitus*) as a manifesta-

¹¹ See, e.g., Camic's discussions of W. I. Thomas, Florian Znaniecki, Robert Park, and Talcott Parsons, among others, in Camic (1986, pp. 1072–75). Camic adds that the historical reasons for this tendency are twofold: on the one hand, the emergence during the late 19th century of Darwinian evolutionary theory and of experimental physiology and, on the other hand, the rise during that same period of a "militantly scientific" new field of psychology. Between them, these developments led to an identification of habitual action with the most elementary behavioral processes of the human organism, akin to those of the lower species (Camic 1986, pp. 1048–49).

tion of human agency.¹² In "The Treatise on Habits," Aquinas (1948, pp. 822, 824) follows Aristotle in associating the habitus with moral virtue: "Virtue is a habitus which is always for good. . . . [It] is a habitus by which a person acts well."

Nineteenth- and 20th-century social thought.—Dewey (1922) contributes to this perspective on habit in *Human Nature and Conduct*, where he describes habits as "active means, means that project themselves, energetic and dominating ways of acting. . . . Habit means special sensitiveness or accessibility to certain classes of stimuli, standing predilections and aversions, rather than bare recurrence of specific acts. It means will" (Dewey 1922, pp. 26, 40–41). Habit emerges as something inherently plastic and educable, rather than a matter of mere stimulus and response. This critique of behavioral reductionism allows Dewey to elaborate the social and psychological foundations for a democratic politics, the goal of which should be to replace the unreflective habits with "intelligent" ones "which experience has shown to make us sensitive, generous, imaginative, [and] impartial" (Dewey 1922, p. 194).

During the mid 20th century, phenomenologists such as Maurice Merleau-Ponty and Alfred Schutz further develop such views, reconceptualizing habit as a form of "prereflective intentionality" (Kestenbaum 1977). For Merleau-Ponty, intentionality is located prior to language in the sedimentation of meaning in the body; the body is conceived of as an "intentional arc" directed toward the world, the vehicle by means of which communication with the world is carried out (Merleau-Ponty 1964, pp. 67; see also Wacquant 1992a). Schutz, on the other hand, emphasizes the social (rather than the embodied) dimension of the prereflexive life world, finding in Weberian ideal-types a model for the schemas and typifications that guide social actors during their routinized daily lives. These typifications provide for the continuity of social knowledge over time; while such knowledge is taken for granted, it nevertheless has a "highly socialized structure" (Schutz 1962, p. 75). This focus upon the routinized prereflexive character of the social world also provides the basis for Harold Garfinkel's ethnomethodology (1984), as well as for the social constructivism of Peter Berger and Thomas Luckmann (1966).

Theories of practice.—In the present day, so-called theorists of practice (Ortner 1984) such as Bourdieu (1977, 1984, 1990; Bourdieu and Wac-

¹² "For Aquinas, . . . a *habitus* puts one's activity more under one's control than it might otherwise be. In this sense, to have a *habitus* is to be disposed to some activity or other—not because one tends to that activity on every possible occasion, but because one finds it natural, readily coped with, an obvious activity to engage in, and so on" (Davies 1992, pp. 225–26; emphasis in the original).

quant 1992) and Giddens (1979, 1984) build upon the insights of both pragmatism and phenomenology, as well as upon earlier traditions of thought. Bourdieu uses the Aristotelian/Thomistic idea of *habitus* to illuminate the formative influences of the past upon the cognitive, corporeal, and intentional structures of empirical action. Through the incorporation of past experiences in the body, he maintains that social actors develop a set of preconscious expectations about the future that are typically inarticulate, naturalized, and taken for granted but nevertheless strategically mobilized in accordance with the contingencies of particular empirical situations. Bourdieu recognizes the compatibility of such notions with the insights of both Dewey and the phenomenologists: "The theory of practical sense presents many similarities with theories, such as Dewey's, that grant a central role to the notion of habit, understood as an active and creative relation to the world" (Bourdieu and Wacquant 1992, p. 122).

In similar fashion, Giddens conceptualizes the agentic dimension of routine behavior in terms of what he calls the *stratification model of action* (Giddens 1979, p. 56). By distinguishing between three levels of consciousness—the unconscious, practical consciousness, and discursive consciousness—he in effect constructs a continuum between the unreflective and reflective dimensions of action. But despite this nod toward discursivity, Giddens gives routinized practical consciousness a privileged place in the explanation of social reproduction, calling routinization the master key of his theory of structuration. Such consciousness emerges out of a background of "tacitly employed mutual knowledge" (Giddens 1979, p. 58), by means of which social interactions are reflexively monitored. In underscoring the agentic moment in the reproduction of structures, he also develops the important idea of *recursivity*: structures (which Giddens defines as "rules and resources") are really only "virtual" structures (paradigmatic patterns) that must be recursively activated within social practices. The agentic dimension of routinized action lies precisely in the recursive implementation of structures by human actors.¹⁴

The Internal Structure of Iteration

We can see that according to many major theorists, habitual and routinized activities are not devoid of agency. Here we elaborate upon these theorists' insights by examining in more detail how agency works to repro-

¹⁴ Giddens (1991) is particularly interested in the concept of routinization because of his ontological presuppositions: he emphasizes the need for "basic trust" and "ontological security" that drives humans to routinize their practices and to give order and stability to their relationships, especially in the face of the growing complexity and diversity of modern society (for a similar perspective, see White [1992]).

duce past patterns of action. For the sake of greater specificity, we subdivide the iterational moment into a number of interrelated components (keeping in mind that these blend into one another in practice); each involves the engagement of a specific kind of schematizing process. Recalling the imagery of the internal chordal structure, we show how this primary orientation toward the past involves different processes of selective recall from past experience, which we distinguish here as *selective attention*, *recognition of types*, and *categorical location*. In addition, we show how these elements shade over into projective and practical-evaluative dimensions of agency. The future and the present now emerge as secondary tones in the chordal composition: the future through *expectation*, the memory-sustained anticipation that past patterns of experience will repeat themselves in successive interactions, allowing relationships to be sustained and reproduced over time, and the present through *maneuver*, the improvisational orientation toward habitual practices, largely tacit and unreflective, which takes place in ongoing dialogue with situational contingencies.

Selective attention.—At any given point in the flow of transactions, social actors are able to focus attention upon only a small area of reality. As Schutz (1964, p. 283) tells us, "There is a small kernel of knowledge that is clear, distinct, and consistent in itself. This kernel is surrounded by zones of various gradations of vagueness, obscurity, and ambiguity." The quality of attention directed at any element or "zone" of knowledge is conditioned by what Schutz calls "systems of relevances," developed over the course of biographical histories and past collective experience, which alert actors to elements of emerging situations that require attention and response. The same idea is expressed in the psychological notion of *gestalt*, which shows how the activity of directing attention is also linked to unconscious processes. Many elements of practical day-to-day activity may require only marginal clarity of consciousness; yet even the semiobscure zone of habitual taken-for-granted activity requires a selective focusing of attention in order to single out the elements of response required to sustain a particular form of interaction.

Recognition of types.—Having directed attention, actors must identify typical patterns of experience and predict their recurrence in the future; to do this, they routinely construct simplifying models by means of which they characterize recurrent aspects of persons, relationships, contexts, or events. As Schutz (1967) puts it, this process of "typification" takes place through a "synthesis of recognition" by which actors recognize the "sameness," "likeness," or "analogy" of an emerging experience with those of the past, either within the actor's direct memory or within a social memory as objectified in various media of communication (see also Alexander 1988, pp. 301–33). While emergent situations never completely match

these simplifying idealizations, actors tend to retrospectively assimilate new experiences to the old by means of an "enveloping" procedure by which differences or faulty "fits" are smoothed over through use of what Garfinkel (1984) calls the *et cetera clause*. Through this active process of recognition and assimilation, actors contribute to a sense of continuity and order within temporally evolving experiences.

Categorical location.—Social actors not only identify similarities between past and present types of experiences; they also locate these typifications in relation to other persons, contexts, or events within matrices composed of socially recognized categories of identity and value. These matrices may be built upon sets of binary oppositions (Lévi-Strauss 1966; Douglas 1985; Bourdieu 1977; Alexander 1988b), which delineate physical, social, and normative categories; as Bourdieu argues, such homologous systems of oppositions constitute transposable schemas by means of which fields of social relationships can be objectively mapped. On the other hand, these classificatory schemas may also be nonbinary and composed of more complex multivalent networks of relationships, containing nuanced lines of inclusion and exclusion, acceptability and nonacceptability, within crosscutting contexts of action. Although for the most part these matrices are unreflective and taken for granted, actors must still exercise effort in order to locate correctly where given experiences fit within them and thus keep social relationships working along established lines.

Maneuver among repertoires.—As we have seen, the employment of routines is not mechanically or situationally determined; rather, it requires a process of selection from practical repertoires of habitual activity. While repertoires are limited by individual and collective histories and may be more or less extensive and flexible, they do require a certain degree of maneuverability in order to assure the appropriateness of the response to the situation at hand. (Here the iterational dimension most closely resembles what we shall later describe as practical evaluation.) In unproblematic situations, this maneuvering is semiconscious or taken for granted, the result of an incorporation of schemas of action into one's embodied practical activity. On the other hand, the application of such repertoires remains intentional insofar as it allows one to get things done through habitual interactions or negotiations (allowing Bourdieu to speak of the paradox of "intentionless intentions"). As Bourdieu notes, there may be much ingenuity and resourcefulness to the selection of responses from practical repertoires, even when this contributes to the reproduction of a given structure of social relationships.

Expectation maintenance.—One of the results of the various forms of schematization described above is that they provide actors with more or less reliable knowledge of social relationships, which allows them to pre-

dict what will happen in the future. These patterns of expectations give stability and continuity to action, the sense that "I can do it again," as well as "trust" that others will also act in predictable ways (Schutz 1967; Garfinkel 1963, 1984). (Here we encounter the subtone in the chordal structure of iteration that most approximates the projective dimension of agency.) The maintenance of expectations regarding how oneself and others will act is not an automatic process: one's expectations about the future can break down (requiring what Garfinkel calls *repair*) due to disruptions, misunderstandings, and changes in systems of relevance. The maintenance work that goes into sustaining expectations has practical as well as ontological importance, allowing not only for a sense of consistent identity amidst change (Pizzorno 1986; Melucci 1994), but also for social coordination within contingent and interdependent environments.

Iteration in Empirical Research

The iterational orientation of agency has already proved a rich source of research questions in a variety of social science disciplines. Here we explore how such research opens up a number of intriguing lines of inquiry into the reciprocal relationship—the ongoing dialogue or conversation—between the agency in its iterational modality and a wide range of temporal-relational contexts of action.

Cultural competences.—Research building upon Bourdieu's notion of habitus proves highly useful in showing how different formative experiences, such as those influenced by gender, race, ethnicity, or class backgrounds, deeply shape the web of cognitive, affective, and bodily schemas through which actors come to know how to act in particular social worlds. Ann Swidler (1986) evokes Bourdieu in speaking of the "cultural toolbox" of practical competences that predispose actors to feel a fit within some actions and not others. Although Loïc Wacquant (1992*b*) criticizes the implicit instrumentalism of Swidler's account, his work on boxing in Chicago ghetto neighborhoods sounds similar themes by exploring how embodied competences and classificatory schemas first learned within the street environment underlie boxers' subsequent engagement of the "pugilistic field." Likewise, Michèle Lamont's (1992) research into money, morals, and manners in France and the United States examines how classificatory schemas developed within particular class, race, and national settings influence the boundary work of social actors in articulating tastes and aspirations, as well as in distinguishing themselves from other social groups. In such ways, these writers claim, the agentic reactivation of schemas inculcated through past experience tends to correspond to (and thus to reproduce) societal patterns: "Social structures and cognitive structures

are recursively and structurally linked, and the correspondence that obtains between them provides one of the most solid props of social domination" (Wacquant 1992a, p. 14).

Reproduction through creativity.—While the above authors tend to focus upon the "closeness of fit" between the habitus and subsequent agentic activity, others operating in a similar tradition emphasize the conflictual and contradictory relationships between human agency and social reproduction. For example, Paul Willis (1977), in his study of the cultural creativity of rebellious working-class lads, argues that their interactively generated criticism and rejection of middle-class trajectories was shaped by their working-class experience and leads, ironically, to the reproduction of their subordinate class position. From a social-psychological perspective, William Corsaro demonstrates how children reproduce adult culture through the creative and interactive elaboration of peer routines: "Socialization is not something that happens to children; it is a process in which children, in interaction with others, produce their own peer culture and eventually come to reproduce, to extend, and to join the adult world" (Corsaro 1992, p. 175). Likewise, Garfinkel (1984) shows in a famous case study how "Agnes," an "intersexed person," deploys tremendous effort and ingenuity in order to negotiate the taken-for-granted dimension of social interactions and thereby to *pass* as a woman according to dominant social norms. While these accounts represent heightened degrees of conscious purpose (Garfinkel), creative embellishment (Corsaro), and/or critical penetration (Willis), and thus brush up against the second and third dimensions of agency, the iterational dimension remains primary, since choices continue to reflect a deeper stratum of culturally and social-psychologically rooted predispositions, thereby contributing to the reproduction of social structures.

Life course development.—Recent research on life course development also inquires into the formative influence of past experiences on agentic processes (Berteaux 1981; Elder 1985, 1994; O'Rand and Kreckler 1990). In the tradition of Thomas and Znaniecki (1918), such research explores the connection between social structures and social-psychological development, as manifested in the life trajectories resulting from particular intersections of biography and history. The implication for agency is that neither social structures nor psychological traits in themselves determine habits of action; rather, actors develop relatively stable patterns of interaction in active response to historical situations. For example, Glenn Elder's (1974) study of cohort effects during the Great Depression demonstrates how family interactions amid periods of economic hardship work to shape emotional and cultural resources and thus to precondition subsequent life careers. Other researchers (Kohli 1986; Meyer 1986) focus upon the institutionalized nature of life course trajectories, which socialize individuals

in relation to prestructured stages and pathways; however, they argue that this does not eliminate the role of agency in the construction of life directions: "The individual life course has to be conceptualized not as a behavioral outcome of macrosocial organizations (or of its interaction with psychological properties of the individual) but as the result of the subject's constructive activity in dealing with the available life course programs" (Kohli 1986, p. 272).

Typification within organizations.—Finally, the importance of habit and routine in shaping interactions is stressed in organizational analysis, particularly by the so-called new institutionalists (Powell and DiMaggio 1991; Meyer and Rowan 1991; Zucker 1977, 1983; March and Olsen 1976, 1984). Reacting against overly instrumental and purposive views of organizational life, many of these researchers draw heavily upon ethnomethodological, phenomenological, and cognitive approaches, emphasizing the routinized, taken-for-granted (or "scripted") quality of knowledge and action that makes organizations relatively stable and resistant to change. Institutional decisions do not develop through rational cost-benefit analysis, but rather are embedded in established routines and become "rationalized" (and thereby legitimated) only through retrospective accounting processes. This approach allows such researchers to argue that the persistence and/or resistance to change of practices within organizations may be due less to social sanctions or to formal structure than to the degree of institutionalization of informal patterns of shared beliefs and socialized expectations (Zucker 1977; Meyer, Scott, and Deal 1981). The strong formative influence of the past can also be seen in the perseverance of organizational procedures even in the face of inefficiency, due to the imprint of founding practices that commit organizations to routines (Nystrom and Starbuck 1984; Powell 1986).

THE PROJECTIVE DIMENSION OF AGENCY

One key limitation of many contemporary theories of agency is that they tend to restrict the discussion of human agency to its iterational dimension. While such theorists as Bourdieu and Giddens do, in fact, recuperate the creative, improvisational, and foresightful dimensions of the implementation of practical schemas of action—what we call here maneuver and expectation—they focus upon a low level of reflectivity and do not show us how such schemas can be challenged, reconsidered, and reformulated.¹⁴ By contrast, we maintain that human actors do not merely repeat

¹⁴ This is not to say, on the other hand, that these authors see change as impossible; Giddens's idea of "discursive consciousness" and Bourdieu's calls for a "reflexive sociology" suggest that each believes a certain increase in freedom and flexibility of action is possible, as one becomes more conscious of one's situation. However, their frame-

past routines; they are also the inventors of new possibilities for thought and action (see also Joas 1993). To understand this creative reconstructive dimension of agency, we must shift our analytic attention away from actors' orientation toward the past and focus upon how agentic processes give shape and direction to *future* possibilities. We argue that an imaginative engagement of the future is also a crucial component of the effort of human actors. As they respond to the challenges and uncertainties of social life, actors are capable of distancing themselves (at least in partial exploratory ways) from the schemas, habits, and traditions that constrain social identities and institutions. This capacity for what Mead calls "distance experience" enables them to reconstruct and innovate upon those traditions in accordance with evolving desires and purposes. The subset of words used to describe this ability has ranged from the strongly purposive terminology of goals, plans, and objectives to the more ephemeral language of dreams, wishes, desires, anxieties, hopes, fears, and aspirations. In this article, we term it the *projective* dimension of human agency.

In our view, projectivity is neither radically voluntarist nor narrowly instrumentalist; the formation of projects is always an interactive, culturally embedded process by which social actors negotiate their paths toward the future, receiving their driving impetus from the conflicts and challenges of social life. The locus of agency here lies in the *hypothesization* of experience, as actors attempt to reconfigure received schemas by generating alternative possible responses to the problematic situations they confront in their lives. Immersed in a temporal flow, they move "beyond themselves" into the future and construct changing images of where they think they are going, where they want to go, and how they can get there from where they are at present. Such images can be conceived of with varying degrees of clarity and detail and extend with greater or lesser reach into the future; they entail proposed interventions at diverse and intersecting levels of social life. Projectivity is thus located in a critical mediating juncture between the iterational and practical-evaluative aspects of agency. It involves a first step toward reflectivity, as the response of a desirous imagination to problems that cannot satisfactorily be resolved by the taken-for-granted habits of thought and action that characterize the background structure of the social world.¹⁵

works do not help us to analyze this possibility, nor do they give us the tools to recognize it in the course of doing empirical research.

¹⁵ Here we need to take great care to avoid misinterpreting what we call the future-oriented aspect of imagination. The desirous imagination is certainly directed toward the past as well as the future; the reconstructive dimension of memory has been well documented by research in this area (Hobsbawn and Ranger 1983; Schwartz 1991; Halbwachs 1992; Olick and Levy 1997, Olick 1997). Mead himself (1932, p. 12) makes this point by insisting that "the past (or the meaningful structure of the past) is as hypothetical as the future." He also stresses, however, that the reason actors engage

Projectivity: The History of a Concept

We wish to stress from the outset that projectivity does not always generate morally superior or desirable engagements with problematic situations. Its potential inventiveness can yield responses as benign and mundane as the projects to grow a garden, to start a business, or to patch up a family relationship, or as sweeping and destructive as the project to establish a 1,000-year Reich. We also wish to stress that not all time periods, cultures, theoretical traditions, or even individuals are equally projective. As Niklas Luhmann (1990) points out, "ancient" conceptions of time (according to which an "enduring present" confronts a temporal flow in which the future is largely predetermined by the past), can be clearly distinguished from "modern" conceptions, in which experience is conceived of as moving toward an indeterminate future, which is purposefully constructed through means-ends rationality. Moreover, many non-Western cultures have alternative constructions of the relationship between past, present, and future, which constrain and enable particular forms of social creativity and reproduction. Our premise is simply that the specific culturally embedded ways in which people imagine, talk about, negotiate, and make commitments to their futures influence their degree of freedom and maneuverability in relation to existing structures (i.e., it *matters* to what degree they understand time as something fixed and determinate, or conversely, as something open and negotiable). These points will become clearer as we examine the historical development of the notion of projectivity in philosophical thought.

Classical and Enlightenment conceptions.—From the Hebraic and ancient Greek traditions, we gain important early conceptions of the projective capacity of human beings. In *Exodus and Revolution* (1985), Michael Walzer offers a compelling interpretation of early biblical narratives, showing how visions by the Jewish people of the future and their own relationship to it—ideas of the covenant, redemption, and promised land—came later to influence Christian narratives of redemption, as well as the discourse of revolutionary politics in the modern world. Within the

in such imaginative reconstruction of past experience is that they confront emergent situations involving new future horizons, that is, the reconstruction of the past is carried out with (more or less explicit) reference to future desires, concerns, and possibilities. We can make the even stronger claim that as action within a given context becomes more self-reflective, the future dimension gains in salience; this is implied, as Joas (1985, p. 192) points out, by Mead's insistence that all self-reflective activity, regardless of the richness with which it engages the past, is "essentially referred to the future . . . It directs itself to the organism's present attitudes that have been formed by the past, becomes aware of their implicit reference to the future, and thereby becomes capable of experimentally testing alternative future possibilities in the present and then deliberately to construct the plan of its own action."

more static destiny-bound framework of the ancient Greeks, however, the future did not have the centrality it has today as an object of human imagination and action (Kearney 1988). Plato was deeply suspicious of the imagination as a source of illusion, irrationality, and immorality, in opposition to the pure, ideal, and eternal world of rational form. From Aristotle's realist epistemology, on the other hand, came the beginnings of a more benign view of the imagination as a psychological link between sensation and reason, which, while not exactly "productive" in the way Kant and the later Romantics would see it, did provide the basis for rational deliberation about the future by allowing social actors to transcend the bounds of sensible experience. Aristotle also gave us the key conception of the telos of action as a basis for means-ends rationality, a view that provides philosophic grounding for prevailing Western instrumentalist narratives about the future.

Tensions between these two contributions of the Aristotelian legacy can later be found in early modern divisions between an affirmation of the moral conscience and the transcendental imagination (which is idealized as the "privileged expression of human freedom" [Kearney 1988, p. 175]), and the abstractly rational—and imaginatively impoverished—instrumentality of the utilitarian tradition. These conflicting concerns eventually gained systematic expression in the dualist philosophy of Kant (which accorded primacy, however, finally to the "practical" or transcendental moment), as well as in the Utilitarian and Romantic currents of the late 18th and 19th centuries. They also gained expression in the Hegelian and Marxist traditions, with their focus upon the telos of history and the relation between objective interests and subjective liberation (see Marx 1978). As we have seen, in present-day sociology, these currents most strongly manifest themselves, on the one hand, in rational choice perspectives, with their stress upon purposive-rational action and, on the other hand, in normative approaches that stress cultural ideals and moral action, the pursuit and realization of ultimate values.

Phenomenological and existentialist perspectives.—Beginning in the late 19th century, we encounter yet another line of reasoning—that of phenomenology and existentialism—that contributes to the development of theories regarding the projective dimension of agency. Building upon Edmund Husserl's (1960) theory of the temporal structure of experience, as well as the passionate dialectics of Søren Kierkegaard, philosophers in this tradition depict actors as "thrown" into historically evolving situations; out of the anguish, uncertainty, and longing that arise from the condition of "becoming," actors necessarily "project" themselves into their own possibilities of being. Reflection about the future is characterized by emotional engagement, "for when existence is interpenetrated with reflection it generates passion" (Kierkegaard 1944, p. 313). Martin Heidegger

(1962) terms this dimension *care* (*Sorge*), the preconscious affective engagement of the world that constitutes the "forestructure" of action; actors invest effort in the formulation of projects because in some way or other they *care* about (not just "have an interest in") what will happen to them in the future.¹⁶ As Jean-Paul Sartre (1956) later stresses, this emotional engagement of the future always implies a thrust to surpass our basic condition of incompleteness: "The fundamental project of the for-itself is to achieve a coincidence with what it lacks" (Bernstein 1971, p. 139).

The bridge from the existential and phenomenological traditions to the sociological preoccupation with shared meaning is made by Schutz (1962, 1967), who stresses "the project" as a fundamental unit of action. Schutz (1967, p. 61) brings Husserl's basically epistemological observations into the realm of action theory, pointing out that "*the meaning of any action is its corresponding projected act.*" Projects represent the completed act-to-be as imagined in the future perfect tense; "*The unity of the action is a function of the span or breadth of the project*" (Schutz 1967, p. 62; emphasis in the original). Here Schutz takes up the question bracketed by rational choice theory: he is interested not in behavioral outcomes, but rather in how forward-looking (but not always utility-maximizing) actors actually construct choices out of fluid and shifting fields of possibilities. For Schutz, purposeful action is rarely guided by the abstract objective analysis of means and ends, or by the clear choice between alternatives, that rational choice theorists propose (ironically, in common with Parsons [Schutz 1978; Joas 1996]). Not only is action limited and shaped by typifications from past experiences, but, more important, both means and ends are always temporally evolving, multiply inflected, and marked by high degrees of indeterminacy. Plans and purposes undergo a continual process of projective "phantasying," in which "rays of attention" are focused upon a plurality of possible future states until choices detach themselves, "like overripe fruit," from the subjective horizons of future actions (Schutz 1967, pp. 67–68).¹⁷

Pragmatist perspectives.—While the existentialist and phenomenological traditions highlight the centrality of projects for human life, they prove

¹⁶ Cornelius Castoriadis (1987, p. 87) draws heavily upon Heidegger—as well as Aristotle and Marx—in his own theory of "the imaginative constitution of society": "To do something, to do a book, to make a child, a revolution, or just doing as such, is projecting oneself into a future situation which is opened up on all sides to the unknown."

¹⁷ In contrast to most rational choice theorists, Schutz (1967, p. 69) maintains that choices are highly unstable and only gain relative clarity after the act has been completed, through *ex post facto* reflection: "The error is to suppose that the conscious state, which only exists after the deed is done, lies back at some 'point of duration' before the actual choice."

less helpful in showing what projects are good for—that is, how our projective capacity is essential to problem solving within a community. Here, once again, we can turn to the pragmatists, who in addition to their concern with routine, are deeply attuned to the imaginative flexibility of actors' deliberations about the future. Dewey (1981, p. 61), for example, characterizes the experimental relationship with the future as an essential dimension of human action: "Experience in its vital form is experimental, an effort to change the given; it is characterized by projection, by reaching forward into the unknown; connection with the future is its salient trait." Human intelligence is based upon the capacity to "read future results in present on-goings" (Dewey 1981, p. 69); this projective capacity permits the kind of responsive choice and inventive manipulation of the physical and social worlds that is so essential to democratic participation. Likewise, Mead (1934) stresses the essentially intersubjective dimension of projectivity, arguing that our basic self-concept is developed from the capacity to project ourselves into the experiences of others. The imaginative capacity of the "I" to move between multiple situationally variable "me's" is what constitutes freedom and maneuverability in relation to established roles, as well as making possible social coordination, joint problem solving, and collective projects of social reform. In the pragmatist view, projects are not constituted merely by "thrownness" into an uncertain world that condemns us to freedom, but also by the practical exercise of that freedom along with others in pursuit of a common good.

The Internal Structure of Projectivity

As the foregoing discussion suggests, the concept of projectivity has a rich legacy in philosophy and in sociological theory. Our own conception builds critically upon the insights of the above-mentioned theorists but seeks to give a more concrete elaboration of how projectivity actually works in social processes. As in the previous section, we outline several important processes involved in the projection of future action, keeping in mind again that these overlap with and feed into one another, interacting in an open-ended, recursive, and synergistic fashion. We differentiate between three dominant tones within the internal chordal structure of projectivity: *narrative construction*, *symbolic recomposition*, and *hypothetical resolution*. In addition, we again point to secondary tones that orient actors toward the other two dimensions of time: relationships to the past through a retrospective-prospective process of *identification*, in which possible trajectories are located against a backdrop of prior typifications from experience, and relationships to the present through *experimentation*, in which alternative courses of action are tentatively enacted in response to currently emerging situations.

Anticipatory identification.—Alternatives are seldom clearly and neatly presented, but neither is the future an open book. Understanding the limited and yet flexible structure of future possibilities involves the work of identifying patterns of possible developments in an often vague and indeterminate future horizon. As Schutz (1967) tells us, this anticipatory work is done by means of a retrospective engagement with one's prior "stock of knowledge" as stored in typifications, repertoires, and social narratives. This retrospective-prospective process shows the essential role of memory in the mapping of future trajectories of action. (In this way, it draws the past into the internal structure of projectivity.) We draw upon past experiences in order to clarify motives, goals, and intentions, to locate possible future constraints, and to identify morally and practically appropriate courses of action. Such anticipatory identifications are never accomplished once and for all, but rather are subject to continual reevaluation in light of the shifting and multidimensional character of human motivations and social relationships.

Narrative construction.—Such identification of typical trajectories is closely tied to the construction of narratives that locate future possibilities in relation to more or less coherent causal and temporal sequences. While narratives are not identical with projects (since narratives represent a particular cultural structure that may exist independently of intentionality), they do provide cultural resources by which actors can develop a sense of movement forward in time (i.e., the proverbial beginning, middle, and end). Jerome Bruner (1986) notes that the plots of such stories contain at least three basic elements: *plight*, *character*, and *consciousness*; these elements help actors to visualize proposed resolutions to lived conflicts (see also Taylor 1989). All social groups possess repertoires of stories that serve as temporal framing resources and that help to define membership in a community (Carr 1986; Somers 1992); the degree of specificity and complexity with which futures are imagined is closely related to the salience of existing narratives and the "careers" (White 1992) that they present as both morally and practically acceptable. While narratives provide "maps of action" (Ricoeur 1991) and thus help to institutionalize stages in the life course (Meyer 1986), they also, because of their flexible and metaphoric structure, can be used to experimentally posit new resolutions to emerging problems.

Symbolic recomposition.—The projective imagination works in a way analogous to the capacity of metaphor to create semantic innovation; it takes elements of meaning apart in order to bring them back together again in new unexpected combinations. Paul Ricoeur (1991, pp. 173–74) describes the imagination as the "free play of possibilities in a state of non-involvement with respect to the world of perception or of action." Actors playfully insert themselves into a variety of possible trajectories

and spin out alternative means-ends sequences, thereby expanding their flexible response to a given field of action. In this play of scenarios, (relatively) freed of practical constraints, symbolic codes, schemas, and narratives can be creatively reconfigured due to their multivocal, homologous, and transposable character (Alexander 1988, pp. 301–33). This process has an intersubjective transactional dimension; for example, in game theory, actors make decisions on the basis of imaginative scenarios regarding the simultaneous imaginative projections by other actors (Axelrod 1984). In a potentially less agonistic fashion, joint projections of action scenarios provide communicative bases for the formulation of new strategies for collective action (Melucci 1989), as well as for the development of new social policies, normative ideals, or ways of organizing institutions (McLoughlin 1978; Castoriadis 1987).

Hypothetical resolution.—After surveying possible scenarios of action, actors face the task of proposing hypothetical resolutions that will adequately respond to the moral, practical, and emotional concerns arising from lived conflicts. The fact that all of our conflicts are overdetermined and that our sense of relevance changes over the course of a lifetime, usually means that such resolutions will be *synthetic* in nature; they will often attempt to resolve several conflicts simultaneously and to incorporate different fields of intended action. A career project, for example, may jointly address a person's desire for money, status, accomplishment, and creative expression, as well as the hope to make a difference in the wider world. Likewise, by participating in social movements, one may attempt to resolve social problems while simultaneously gaining the opportunity for peer recognition, solidarity, rebelliousness, and organizational achievement. While all of these resolutions are not necessarily present at the outset as clearly articulated goals of action (and may be understood, if at all, only through *ex post facto* reflection), most actors, when pressed, will give more or less differentiated and multivalent descriptions of what they "want" or "intend" in their plans to pursue a particular course of action.

Experimental enactment.—This final dimension of projectivity rests on the borderline between imagination and action (and hence between the future and the present); once scenarios have been examined and solutions proposed, these hypothetical resolutions may be put to the test in tentative or exploratory social interactions. Psychologists such as Erik Erikson (1968) speak of this as "role experimentation," particularly salient during adolescence, by means of which individuals try out possible identities without committing themselves to the full responsibilities involved. Experimental enactments often have ritual overtones, which have been studied in versions of symbolic interactionism (Goffman 1959) as well as dra-

maturgic anthropology. Victor Turner (1974), for example, describes the "social dramas" that are enacted during "liminal periods" in which societies ritualistically reverse social roles. Although Turner stresses how such dramas reinforce the social order, we would argue that these liminal experimental periods may also have a transformative and renovational effect upon the larger culture, as new possibilities for human interactions are imagined, tested, and (perhaps) defined on a collective scale.

Projectivity in Empirical Research

In considering how past patterns of interaction are imaginatively recomposed to generate new future possibilities, we open up a richly suggestive field for sociological research. This is in contrast with much of empirical sociology, where, despite its extensive philosophical legacy, the notion of projects has largely been ignored, due in part to its perceived subjective nature and the apparent incompatibility of "imaginative" phenomena with behavioral observation, survey techniques, and macrostructural analysis. We argue that projectivity needs to be rescued from the subjectivist ghetto and put to use in empirical research as an essential element in understanding processes of social reproduction and change. Many of the elements outlined above have, in fact, been addressed by a wide body of literature in various social science disciplines, albeit in an undertheorized and residual way. Here we discuss several of these approaches (and their limitations), in order to point toward future research on the interplay between the projective dimension of agency and the different temporal-relational contexts of action.

Time perspectives.—While most life course approaches in sociology have tended to focus upon the influence of past experiences on subsequent life paths, a well-developed subfield in social psychology has explored questions more directly linked to projectivity. Since the 1940s, research has been carried out on "time perspective" and its influence on such matters as academic success and civilian morale (Lewin 1948); more recently, researchers in this area have investigated changes in time perspective during different developmental periods, such as childhood, adolescence, middle adulthood, and old age. Of particular relevance to projectivity are studies of the construction of future expectations, examining such factors as variability in the density and extension of imagined future events, linked to cognitive development and/or particular social contexts such as family or class background (Cottle and Klineberg 1974; Devolder and Lens 1982; Wessman and Gorman 1977; Greene 1986, 1990). While much of this research is limited by overly behavioral and correlational assumptions, recent theorists of narrativity have added an interpretive dimension

to life course studies (Gergen and Gergen 1983, 1984, 1988; Bruner 1986; Sarbin 1986), exploring how personal conceptions about past and future are transformed at key moments of transition and/or crisis (Riegel 1975, 1977; Cohler 1982).

Prophetic movements.—A second line of work that directly engages projectivity is the extensive literature on prophetic, utopian, and revolutionary movements. While such literature can be criticized for its overemphasis on cultural (as opposed to social-structural or social-psychological) factors, we argue (along with Desroche [1979] and Ricoeur [1991]) that the projective imagination as expressed in collective ideals and aspirations plays a constitutive, not just an epiphenomenal, role in a wide variety of historical phenomena, ranging from millenarian movements, religious cults, alternative communities, and revolutionary organizations, to more generalized forms of cultural revival. For example, Norman Cohn (1977, pp. 16–17) argues that millenarian projections appearing in Europe during the 11th–16th centuries resulted in very different kinds of movements than the more limited localized peasant or artisan revolts of the period: “The usual desire of the poor to improve the material conditions of their lives became transfused with phantasies of a world reborn into innocence through a final, apocalyptic massacre.” Likewise, William McLoughlin (1978, p. 2) claims that major “great awakenings” during periods of uncertainty and change in American history led to cathartic revivals that “eventuated in basic restructurings of our institutions and redefinitions of our social goals.” Finally, both Marxist and non-Marxist historians of revolutions (e.g., Thompson 1966, 1993; Walzer 1965, 1980) explore the projective dimensions of revolutionary movements, which Walzer defines as “conscious attempts to establish a new moral and material world and to impose, or evoke, radically new patterns of day-to-day conduct.” Revolution itself, he concludes, “is a project” (Walzer 1980, pp. 202–3).

Framing processes.—The projective imagination is also a factor in less apocalyptic forms of social movements and efforts at institutional reform. Most work in this well-researched area fails to adequately theorize the projective dimension, due in part to the paradigmatic split during the 1970s and 1980s between “strategy” and “identity” (Cohen 1985). This split, which goes back to the Kantian division between interests and ideals, has had the effect of severing two intrinsically linked dimensions of projectivity: strategies are stripped of meaning and reflexivity, while identities are temporally flattened out and shorn of their orienting power (Mische 1994).¹⁸ Recent attempts to bring the two paradigms together (see Mor-

¹⁸ Mische (1997, pp. 7–8) has further developed this critique in an empirical study of projectivity and social movements, arguing that the concept of projectivity allows

ris and Mueller 1992) have resulted in concepts approximating projectivity, such as that of framing in collective action (Snow et al. 1986; Snow and Benford 1988; Gamson 1992; Tarrow 1992; for an alternative formulation, see Steinberg [1996]). Despite its structuralist overtones, the notion of frames (or more accurately, framing processes) "implies agency and contention at the level of reality construction" (Snow and Benford 1992, p. 136). Framing is both *diagnostic* and *prognostic*: it suggests "a general line of action for ameliorating the problem and the assignment of responsibility for carrying out that action" (Snow and Benford 1992, p. 137). In proposing new social ends as well as different means for arriving at them, actors draw upon—and sometimes extend, rearrange, and transform—the master frames extant in the broader political culture.

Institutional innovation.—A fourth research area in which projectivity is important (but as yet underdeveloped) is that of institutional innovation and change. As we have seen, the new institutionalists reacted against rational-choice views of organizational decision making by, in effect, eclipsing the projective dimension, arguing that institutional purposes are embedded in routines that come to light only in post hoc accounting practices. But recently, some organizational researchers (DiMaggio 1988, 1991; Galaskiewicz 1991; Fligstein 1991; Brint and Karabel 1991) have tried to recuperate the purposeful and conflict-driven aspect of organizations and to pay more attention to processes of institution building and reform. Paul DiMaggio (1991), for example, invokes the language of projectivity (albeit without theorizing it) in his study of the struggles of museum professionals over the model of art museum to be imposed on a developing organizational field. DiMaggio (1991, p. 277) shows how opportunities for "professional projects" "reinforced the awareness that they were part of a collective enterprise, and thus the likelihood that they would look to one another as models and as sources of innovation." These projects were constructed by drawing contentiously upon the "Western cultural account" of justice and progress (Meyer, Boli, and Thomas 1987), showing the importance of narrative reconstruction in the development of collective projects of action.

us to supersede the split between rational choice and norm-based (or identity-based) perspectives on collective action: "Projects are simultaneously moral, practical, and political in scope, weaving together ideals and interests, protest and proposals, utopian alternatives and pragmatic assessments of opportunity structures." Similarly, projectivity challenges the divide between rational choice and cultural determinism: "Projects are the means by which actors imaginatively formulate purposive actions, but these are always composed from the cultural narratives and repertoires at hand. . . . In contrast to the abstract voluntarism of rational-choice theory, moreover, the construction of projects is situationally contingent, subject to learning processes and revision, and always surrounded by a high degree of uncertainty."

THE PRACTICAL-EVALUATIVE DIMENSION OF AGENCY

The final variation we examine in the chordal triad of agency is that which responds to the demands and contingencies of the present. Even relatively unreflective routine dispositions must be adjusted to the exigencies of changing situations; and newly imagined projects must be brought down to earth within real-world circumstances. Moreover, judgments and choices must often be made in the face of considerable ambiguity, uncertainty, and conflict; means and ends sometimes contradict each other, and unintended consequences require changes in strategy and direction. "A rule doesn't [just] apply itself; it has to be applied, and this may involve difficult, finely tuned judgments. . . . There is, as it were, a crucial 'phronetic gap' between the formula and its enactment" (Taylor 1993, p. 57). The problematization of experience in response to emergent situations thus calls for increasingly reflective and interpretive work on the part of social actors. This exercise of situationally based judgment has been variously termed practical wisdom, prudence, art, tact, discretion, application, improvisation, and intelligence; here we designate it as the *practical-evaluative* dimension of agency.

The primary locus of agency in its practical-evaluative dimension lies in the *contextualization* of social experience. Again, we echo the pragmatists in stressing the communicative transactional dimension of such processes; through deliberation with others (or sometimes, self-reflexively, with themselves) about the pragmatic and normative exigencies of lived situations, actors gain in the capacity to make considered decisions that may challenge received patterns of action. This communicative process is what distinguishes the "strong" situational moment of deliberative decision making from the "weak" situatedness of what we call, in the iterational dimension, tacit maneuver. By increasing their capacity for practical evaluation, actors strengthen their ability to exercise agency in a *mediating* fashion, enabling them (at least potentially) to pursue their projects in ways that may challenge and transform the situational contexts of action themselves (although, given the contingency and uncertainty of interactions, the consequences of their actions cannot be controlled and will often "feed back" in ways that necessitate new agentic interventions).

Practical Evaluation: The History of a Concept

Despite its long history, the concept of practical evaluation has received less sustained and systematic treatment during modern times than it did in the ancient or medieval periods. In contemporary action theory and moral philosophy, it has been overshadowed by an emphasis upon clear and explicit rules of conduct, concepts that permit relatively little scope for the exercise of situationally based judgment. In social theory, modern

concerns with explicit decision procedures and a widespread "flight from ambiguity" (Levine 1985) and judgment have become evident in a host of analytical perspectives—not only rational choice theory, but also less explicit yet equally instrumentalist conceptions of social action, dating back at least to Max Weber's discussions of *Zweck-* and *Wertrationalität* (Weber 1978). Even Durkheim (1961, pp. 31, 26) sees morality, by definition, as a "system of commandments," "an infinity of special rules [that are] fixed and specific." "To the extent," he writes, "that the rule leaves us free [and] does not prescribe in detail what we ought to do, the action being left to our own judgment, to that extent there is no moral valuation" (Durkheim 1961, pp. 23–24).

Aristotelian perspectives on practical wisdom.—We must return once again to Aristotle's writings on ethics for one of the earliest (and most fully developed) theories of prudence or practical wisdom. In marked contrast to later rule-based theories, Aristotle holds that "three features of 'the matter of the practical' . . . show why practical choices cannot be adequately and completely captured in a system of universal rules" (Nussbaum 1986, pp. 303–4): the mutability of the particular, its indeterminacy (complexity and contextual variety), and its inherent nonrepeatability.¹⁹ Also, the values, rules, and principles that are constitutive of a good human life are themselves plural and incommensurable; hence, a concern for situated judgments supplants any simple belief in the unproblematic application of universal norms or imperatives (Nussbaum 1986, pp. 303–4, 294–95). In Aristotle's view, practical wisdom can refer variously to means or to ends; it can be either strategic and calculative—in which case, he says, we can speak of persons as being clever, crafty, or cunning—or it can be concerned with broader questions of the good life itself (Aristotle 1985, p. 153). Aristotle sees practical wisdom as intrinsically communicative in nature; that is, it entails a deep involvement and participation in an ongoing community of discourse. Far from being purely individual or monological, it remains open to dialogue and persuasion and is profoundly implicated in common values, interests, and purposes.

Theories of judgment and critical deliberation.—A significant break with this legacy comes about with Kantian ethics, which regards prudence not as a virtue, as did so many earlier moral theories, but rather as a vehicle for cold and selfish calculation, expediency, and pragmatism. And yet, especially in his later work, even Kant indirectly provides a theory of practical evaluation.²⁰ Moreover, he adds that practical judgments (spe-

¹⁹ "Perhaps the most obvious and astonishing absence from Aristotle's thought for any modern reader is that there is relatively little mention of rules anywhere in the *Ethics*" (MacIntyre 1981, p. 141).

²⁰ For Kant's early critical views of prudence, see Kant (1956), pp. 16, 37–38; (1964), p. 83. For his later work on practical judgment, see Kant (1971), pp. 389–90. In the

cifically, judgments of taste) fall within the potentiality of all persons since they "depend . . . on our presupposing the existence of a common sense [*sensus communis*]" (Kant 1951, p. 83). Kant links such "common sense" to what he calls the capacity for an "enlarged mentality," in which judgment is carried out by abstracting from one's own limited experience in order to put oneself in the position of everyone else and thus to deliberate over the collective good. Such an idea recalls Aristotle's notion of a community of discourse, as well as the more distinctively modern theme of autonomy, since judgment no longer depends upon the subjectivity and caprice of concrete individuals.

More recent examples of theories that fully embrace the critical and dialogic aspects of practical evaluation can be found in the writings of John Dewey, Hannah Arendt, and Jürgen Habermas. Dewey subsumes Kant's insights on reflective judgment into his own pragmatist—and eminently relational—theory of judgment. In "The Logic of Judgments of Practice" (1985), he points out that all such judgments begin with a problematic experience, a fork in the road, which they attempt experimentally to resolve. Judgments gain intersubjective validity from assuming the standpoint of a *sensus communis*, "a whole of common interests and purposes" (Dewey 1978, p. 286). Arendt (1984, p. 36) also expands upon Kant by maintaining that reflective judgment is not limited to aesthetics but represents "the most political of man's mental abilities." She builds upon Kant's notion of the enlarged mentality, which she terms "representative thinking," describing it as the ability to see things from the perspective of others, "an anticipated communication with others with whom I know I must finally come to some agreement" (Arendt 1977a, pp. 220; see also Arendt 1977b; Benhabib 1992b). And finally, Habermas (1990, 1993) engages Kant's doctrine of judgment while insisting that he is correcting Kant's ethical rigorism; to a Kantian "discourse of justification" he adds a more Aristotelian "discourse of application." In developing his theory of communicative action, Habermas retains a Kantian emphasis upon deliberation and intersubjective validity, even as he objects to the emptiness of Kantian ethics itself.²¹

Critique of Judgment, Kant (1951, p. 18) distinguishes between "determinate" and "reflective" judgments; the former merely subsume the particular under a rule or universal already given for it, while the latter are "compelled to ascend from the particular in nature to the universal." For Kant, logical and moral judgments belong to the former category, while judgments of taste belong to the latter and necessarily involve practical evaluation.

²¹ A more ambiguous example (from within the Kantian tradition) of implicit reasoning in respect to practical evaluation is Max Weber's (1946) classic discussion of "responsible action" (see also Roth and Schluchter [1979], chap. 2), which requires an "open-eyed" apprehension of concrete situations and of the possible (unintended) consequences of action within them. Weber's analysis is an ambiguous one because, unlike

Feminist theories.—Meanwhile, many feminist thinkers critically draw upon both Aristotelian and Kantian outlooks on practical evaluation in analyzing the particular capacities, experiences, and histories of women, while also generalizing from these experiences to develop broader (less “essentializing”) theories of moral and practical reasoning. One important contribution is Carol Gilligan’s (1982, p. 22) *In a Different Voice*, which stresses gender differences in the use of situated reasoning and a “contextual mode of judgment” and thereby seeks to overcome the limitations of Kant’s abstract universalistic conceptions of moral judgment and action (e.g., Kohlberg 1981). From a very different perspective, Donna Haraway (1988), too, criticizes established understandings of “objectivity,” and calls instead for “situated knowledges” grounded in the particularities of partial “limited” locations. Finally, Seyla Benhabib stresses processes of dialogue and public deliberation in her own communicative conception of practical judgment: there is “no incompatibility,” she writes, “between the exercise of moral intuition guided by an egalitarian and universalist model of moral conversation [Kant] and the exercise of contextual judgment [Aristotle]” (Benhabib 1992a, p. 54; see also Benhabib 1987, 1992c).

The Internal Structure of Practical Evaluation

As the foregoing discussion demonstrates, practical evaluation as a concept is associated with many different forms of activity: with cognitive, moral, and aesthetic judgment as well as with general modes of practical consciousness and action; with expansive ideals of universality, together with more restrictive notions of gendered identities and social positions; with cleverness and calculation, and yet also with enlarged thinking and public deliberation. Here we examine the internal structure of practical evaluation, showing how certain of its dimensions are implicated in all of the manifestations mentioned above. We suggest that three dominant tones within its internal chordal structure can be distinguished as *problem-ization*, *decision*, and *execution*, all of which require the contextualization of projects or of habitual practices within the concrete circumstances of the moment. We also describe two secondary tones: the actor’s relationship to the past is based upon the *characterization* of a given situation against the background of past patterns of experience; and the relationship to the future is characterized by *deliberation* over possible trajectories

those of Dewey, Arendt, and Habermas, it points toward a decisionistic ethics (“Here I stand; I can do no other”) and greatly downplays Kant’s original vision of an “enlarged” or “representative” (Arendt) thinking. Weber fails to theorize the intersubjective processes whereby ultimate ends may themselves be chosen by reflective actors in a wise and prudential fashion.

of action, in which actors consider alternative hypothetical scenarios by critically evaluating the consequences of implementing these within real-world situations.

Problematisation.—The first analytical component of practical evaluation consists in the recognition that the concrete particular situation at hand is somehow ambiguous, unsettled, or unresolved. In the case of projects, this recognition entails the apprehension of present reality as in some degree resistant to their immediate and effortless realization, posing challenges in application or contextualization. In the case of iterational or habitual activity, there is also the problem that no new situation is ever precisely the same as ones that came before; all routine activity faces new contingencies to which certain adjustments have to be made. Hence the critical challenge of “analogical transposition” raised explicitly by William Sewell (1992) and addressed as well by Bourdieu and Giddens. Dewey refers to this problem as the objective “incompleteness” of situations: “This incompleteness is not psychical. Something is ‘there,’ but what is there does not constitute the entire objective situation. . . . The logical implication is that of a subject-matter as yet *unterminated*, unfinished, or not wholly given” (Dewey 1985, p. 15). Something must be done—some practical judgment arrived at—that will render the given situation unproblematic, settled, and resolved.

Characterization.—The problematic circumstances at hand must in turn be related to principles, schemas, or typifications from past experience by which they are *characterized* in some fashion. (This component most deeply implicates the past in the moment of practical evaluation.) Does the situation in question call for the activation of a particular iterational or habitual activity? Does it call for the performance of a specific duty, or present itself as a context in which the pursuit of a particular project of action is appropriate or even possible? Speaking in specific reference to moral situations, Benhabib (1992*b*) terms this the problem of “epistemic identification” (while Aristotle calls it “perception” or “understanding,” and Kant discusses it under the rubric of “reflective judgment”). It requires “responding to nuance and fine shading, adapting [one’s] judgment to the matter at hand in a way that principles [or schemas of action] set up in advance have a hard time doing” (Nussbaum 1986, p. 301). Judgments of this nature are emotional (or “passional”) as well as cognitive: “Perception is a complex response of the entire personality” (Nussbaum 1986, p. 309), in which emotions can be seen (with Aristotle) as themselves “intelligent,” educable, and inseparable from intellectual life.

Deliberation.—Plausible choices must be weighed in the light of practical perceptions and understandings against the backdrop of broader fields of possibilities and aspirations. (Here the element of projectivity enters into processes of practical evaluation.) Deliberation involves more than

an unreflective adjustment of habitual patterns of action to the concrete demands of the present; it also entails (at least potentially) a conscious searching consideration of how best to respond to situational contingencies in light of broader goals and projects. Such consideration can take place individualistically or discursively, monologically or within public spaces, recalling the Kantian ideal of an "enlarged mentality." While often employing strategic reasoning or means-ends rationality, it can also require attention to "what *conduces* to the end" (Aristotle 1985, p. 63; emphasis added); it therefore entails further specification of habits and projects as well as determination of the specific means for actualizing them. Deliberation applies to conflict among alternative possible ends, no less than it does to the contextualization of singular ends, involving a search for the proper course of action to follow under ambiguous circumstances (Taylor 1985). Finally, deliberation also entails emotional engagement with the particularities of situations; it stands "on the borderline between the intellectual and the passional, partaking of both natures: it can be described as either desiderative deliberation or deliberative desire" (Nussbaum 1986, pp. 307–8).

Decision.—Deliberation aims toward decision (or choice), the resolution to act here and now in a particular way. In certain cases, such resolution entails a highly discrete or circumscribed choice: an actor "finally arrives at a decision." In other cases, it blends indiscriminately into the flow of practical activity, and is only clearly perceived after the fact. In all of these cases, it points in the direction of action within the circumstances of the present and yields a resolution to translate engagement with such circumstances (however passional or implicit) into concrete, empirical intervention. It should be noted that not all choices reflect unambiguous strategies; for this reason, Dewey (1940) speaks of flexible "ends-in-view" rather than of clear and fixed objectives. Certain decisions are provisional, tenuous, and opportunistic, as we shall see below; they may also engage (in a synthetic or polysemous manner) with more than one problematic situation simultaneously. Nor do all decisions lend themselves to easy formulation and explication. Choices can be a matter of tacit adjustment or adaptation to changing contingencies—including feedbacks from experience—as well as the product of articulable explicit reasoning.

Execution.—If deliberation entails consideration or planning, and decision marks a movement toward concrete action, then executive capacity is that capacity "to do the things that tend towards the mark that we have set before ourselves" (Aristotle 1985, p. 169). It is a capacity to act rightly and effectively within particular concrete life circumstances. Ideally, one not only grasps what one ought to do but also how best to set about it in the case at hand. To respond "at the right times, with reference to the right objects, toward the right people, with the right aim, and in the right

way, is what is appropriate and best, and this [is what] is characteristic of excellence" (Aristotle 1985, p. 44). Sometimes even judicious execution, however, entails tragic loss, as when the fulfillment of a duty or realization of a particular vision of the good requires the sacrifice of an equally compelling duty or good (Hook 1974). Execution, in such cases, marks not a happy resolution but rather the fulfillment of a lesser evil. Moreover, even relatively unproblematic instances of execution often create new problems for action further down the road; feedback effects may be initiated over which actors themselves have little control and which they may not even intend. In any case, with execution or action, the arc of practical evaluation is complete: not only deliberation and judgment, but execution as well is required for the contextualization of our habits, ends, duties, and projects.

Practical Evaluation in Empirical Research

Finally, we outline research findings that pertain to empirical manifestations of practical evaluation, in order to convey a clearer sense of what is entailed by this analytical aspect of agency and to show how it can be investigated sociologically. These findings underscore possible ways in which practical evaluation might be elicited in particular contexts and in which it affects in turn the ability of actors to engage with, respond to, and potentially transform their structural environments.

Temporal improvisation.—One research area that provides insight into the temporal contextualization of both ritual and purposive action includes studies of sequencing processes in social interactions. For example, Bourdieu's investigations of the manipulation of the temporal structure of gift exchange reveal that the same gift-giving act can have different meanings at different times, altering the effectiveness of the intended act. Temporal strategies that enable actors to control intervals between expected ritual transactions—for example, by "holding back or putting off, maintaining suspense or expectation," or otherwise manipulating the "tempos" of action—allow them to gain significant material and/or symbolic advantages vis-à-vis their partners in exchange (Bourdieu 1977, p. 71). Additional examples of temporal improvisation include "turn-taking" patterns in everyday conversational interactions. Conversation analysts in the tradition of Schutz and Garfinkel (e.g., Sacks, Schegloff, and Jefferson 1974) investigate the subtleties of timing and delay in the social organization of talk, showing at a micro level how agentic manipulations of time allow actors to engage in repair work, to avoid or (alternatively) initiate conflict, and in myriad other ways to advance their own interests.

Resistance, subversion, and contention.—Another opening for practical

judgment can be seen in the "procedures and ruses" (De Certeau 1984) by which actors can resist and subvert the logics and practices of the established order. Such tactics are "always on the watch for opportunities that must be seized 'on the wing'" (De Certeau 1984, pp. xix–xx). They utilize an "art of placing blows," of "getting around the rules of a constraining space" (De Certeau 1984, p. 18). James Scott (1985, 1990) explores the use of such "tactics of resistance" among oppressed groups and individuals; in his studies of Malay villagers, as well as in broader contexts of slavery, serfdom, caste subordination, colonialism, racism, and patriarchal domination, he uncovers strikingly similar patterns of disguised dissent from what he terms (echoing Goffman [1959]) "official" or "public transcripts." In examining more overt instances of resistance and collective action, Charles Tilly (1986, 1994) also underscores the shrewdness, tact, and situational awareness of individuals and groups, even in the implementation of what he calls "repertoires of contention"; they "perform in dramas in which they already know their approximate parts, [but] during which they nevertheless improvise constantly" (Tilly 1994, p. 15).

Local or prudential action.—Yet another window of opportunity for practical evaluation arises in those structural situations in which no clear expectations for action apply in the first place, settings in which, as Eric Leifer (1988, p. 865) puts it, "roles are not 'givens' that constrain interaction, but something that actors must acquire through interaction." These settings (or "pockets") of role ambiguity necessitate what Leifer calls "local action," in which actors in face-to-face competition avoid claiming "global" roles until their partners signal that such roles will be recognized. A powerful illustration is provided by John Padgett and Christopher Ansell's (1993) study of the rise of the Medici in early modern Florence. Padgett and Ansell (1993, p. 1264, n.9) complicate Leifer's model of local action by speaking of "an entire linked ecology of games, each game layered on top of another," rather than of one single, unitary game. But both accounts concur on the importance of "flexible opportunism—maintaining discretionary options across unforeseeable futures in the face of hostile attempts by others to narrow those options" (Padgett and Ansell 1993, p. 1263).

Political decision making.—A highly contextualized analysis of political leadership and decision making can be seen in work of such authors as Alfred Stepan (1978) and Guillermo O'Donnell and Philippe Schmitter (1986) on the breakdown of (and transitions to) democracy. These writers describe open-ended and contingent sequences of action, underscoring the uncertainties and multiple possibilities confronting actors at each stage of complex reversible processes; whether they be "hard liners," "soft liners," oppositional publics, or military men, political leaders require "good . . . judgment to test the limits of a situation" (O'Donnell and Schmitter 1986,

p. 27). More counterfactually, Barrington Moore (1978) analyzes leadership choices in his discussion of "the suppression of historical alternatives" in Germany after World War I, choices that might have led to a more stable regime and thereby avoided the horrors of Nazism. (That is, was a different policy possible? Why was it not attempted? How about alternative tactics, strategy, and timing?) Leon Trotsky's (1980) assessment of the pivotal role that Lenin played in the making of the Russian Revolution is another classic analysis of situationally contingent decision making; more recently, Timothy Garton Ash (1990) has analyzed the decisive yet almost seat-of-the-pants way in which leaders of the Velvet Revolution orchestrated and channeled events in Czechoslovakia during the crucial months of mid 1989.

Deliberation in publics.—One of the most important applications of judgment, and by extension of the capacity for human agency itself, is deliberation over the proper appropriate ends of action—over what conduces to these ends. Empirical studies of civil society (Cohen and Arato 1992) and of "publics in history" (Emirbayer and Sheller 1996; see also Emirbayer 1992a, 1992b) closely examine such agentic processes of "representative thinking" and collective deliberation. For example, Jane Mansbridge's (1983) ethnography of a New England town meeting and of an urban crisis center concludes that citizen boards are most effective when the judgment and experience of members contributes to common problem solving. Seymour Martin Lipset, Martin Trow, and James Coleman examine the internal dynamics of deliberation within participatory workplaces in their classic sociological study, *Union Democracy* (1962). More recently, analysts such as Alain Touraine and his associates (1983), Lawrence Goodwyn (1991), and Roman Laba (1991) have investigated the processes of collective deliberation that prevailed at the grassroots level during the Solidarity movement in Poland. They demonstrate how Polish citizens arrived at judgments regarding the very nature of their movement, its ultimate ends, and even the ideals to which they aspired through democratic discourse, dialogue, and debate within public spaces.

CHALLENGES FOR FUTURE RESEARCH

In this final section we turn to the question of how the three dimensions of agency—iteration, projectivity, and practical evaluation—enter into different and changing relationships with the temporal-relational contexts of action. The challenge here is to analyze the variable nature of the *interplay* between structure and agency, rather than to understand these as either standing in insurmountable opposition, or, as in currently influential theorizations, being "mutually constitutive" in a direct and stable way.

We contend that as actors alter or shift between their agentic orientations, dialogically reconstructing the internal composition of their chordal triad, they may increase or decrease their capacity for invention, choice, and transformative impact in relation to the situational contexts within which they act. Such a conception opens up compelling questions for future research across many different empirical subfields.

Structure, Action, and Agency

A variety of recent attempts to rethink the relationship between structure and agency have argued that the Kantian dichotomy between ideal and material realms—together with parallel distinctions between free will and necessity, voluntarism and determinism—must be replaced by an outlook that regards these elements as reciprocally constituting moments of a unified social process. Seminal work in this area includes Bourdieu's (1977, 1990) attack on the division between subjectivism and objectivism, as well as Giddens's (1979, 1984) theory of structuration, which characterizes structure and agency as mutually constitutive (and hence inseparable) elements. This notion has been a salutary and fruitful one for sociological theory, making possible empirical research that underscores both the causal significance of structure as the constraining and enabling conditions of action, and of praxis as "an active constituting process, accomplished by, and consisting in, the doings of active subjects" (Giddens 1976, p. 121). But it has also brought in its train several theoretical disadvantages. Foremost among these is a tendency toward what Margaret Archer (1982, 1988) terms the "fallacy of central conflation": the tendency to see structure as so closely intertwined with every aspect of practice that "the constituent components [of structure and agency] *cannot* be examined separately. . . . In the absence of any degree of autonomy it becomes impossible to *examine* their *interplay*" (Archer 1988, pp. 77, 80; emphasis in the original).²²

²² Strictly speaking, Archer means by "central conflation" an elision of the two key elements of "Cultural System" and "Sociocultural Interaction." We generalize from her criticisms to make a broader point about the relationship between agency and its plurality of structural contexts. If, as Archer (1988, pp. 89–90) puts it, "the powers of Mephistopheles [structure] ultimately depend on Faustus [agency] continuing to invoke them," the constraining and enabling powers of specific actual structures cannot be determined. And correspondingly, if actors "are assumed to enjoy a constant degree of transformative freedom," then the circumstances under which one encounters "more voluntarism" or "more determinism" also cannot be specified (on this point, see also Alexander [1994]).

What becomes eclipsed in the notion of the inseparability of structure and agency is the degree of changeability or mutability of different actual structures, as well as *the variable (and changing) ways in which social actors relate to them*. In most central-conflationist views, the constitutive relationship between agency and structure is held analytically constant. We argue, by contrast, that while the temporal-relational contexts of action influence and shape agency and are (re)shaped by it in turn, the former is never so deeply intertwined with every aspect of the latter that these different analytical elements cannot be examined independently of one another. The agentic orientations of actors (along with their capacity for inventive or deliberative response) may vary in dialogue with the different situational contexts to which (and by means of which) they respond. While human agency represents the possibility for imaginative distancing from (and communicative evaluation of) received structures, agentic processes themselves assume diverse empirical forms in response to the specific contexts within which action unfolds. We might therefore speak of the *double constitution of agency and structure*: temporal-relational contexts support particular agentic orientations, which in turn constitute different structuring relationships of actors toward their environments. It is the constitution of such orientations within particular structural contexts that gives form to effort and allows actors to assume greater or lesser degrees of transformative leverage in relation to the structuring contexts of action.

Here it is important to be perfectly clear about our analytical distinctions: the foregoing formulations are based upon a threefold differentiation between agency, action, and structure. While what we have called "agentic orientations" vary in their concrete manifestations, agency itself remains a dimension that is present in (but conceptually distinct from) all empirical instances of human action; hence there are no concrete agents, but only actors who engage agentially with their structuring environments. We concur with Alexander (1992, pp. 1–2) that the "identification of actor and agency" renders one "guilty of [the fallacy of] misplaced concreteness. Rather than replacing or reinterpreting the familiar dichotomy between actors and structures, [this] identification . . . actually reproduces it in another form. . . . Actors per se are much more than, and [simultaneously] much less than, 'agents' [alone]." All social action is a concrete synthesis, shaped and conditioned, on the one hand, by the temporal-relational contexts of action and, on the other, by the dynamic element of agency itself. The latter guarantees that empirical social action will never be *completely* determined or structured. On the other hand, there is no hypothetical moment in which agency actually gets "free" of structure; it is not, in other words, some pure Kantian transcendental will.

Empirical Propositions

Given these theoretical formulations, the empirical challenge becomes that of locating, comparing, and predicting the relationship between different kinds of agentic processes and particular structuring contexts of action. Here we take a step beyond important recent initiatives in this direction, such as that of Sewell (1992), which focuses primarily upon the structural side of such variation. While building upon the work of Giddens and Bourdieu, Sewell (1992, p. 16) criticizes the overly reproductive conceptions of these authors, arguing that "a theory of change cannot be built into a theory of structure until we adopt a far more multiple, contingent, and fractured conception of society—and of structure." Agency, in his view, consists primarily in the capacity of resource-equipped actors to act creatively through the transposition of existing schemas into new contexts. He notes that "agency differs enormously in both kind and extent," but attributes this difference primarily to the "nature of the particular structures that inform those social worlds" (Sewell 1992, pp. 20–21). While this framework allows Sewell to advance several highly suggestive propositions regarding different rates of change among such structures as language, states, and capitalist economies, he fails to offer any theorization of differences in agentic capacity that are not inseparably bound to structural qualities. Moreover, he does not examine the internal composition of agency itself, and, in particular, the temporal orientations of agency that we have discussed in this article.

Where might we look for evidence of such variation in agentic capacity? How might we locate what we have called the *double constitution* of agency and structure (i.e., how temporal-relational contexts constitute the patterns of response that shape agentic orientations, which go on to constitute different mediating relationships of actors toward those contexts)? Building upon Mead's suggestion that it is the sociality of experience that drives the development of agentic capacities, we offer three lines of questioning through which these analytical formulations might point to new initiatives in empirical research.

1. *How do different temporal-relational contexts support (or conduce to) particular agentic orientations?* This initial question might be considered the first constitutive dimension of the study of agency, in which agentic orientations are held steady in order to examine the formative influences upon them of different kinds of situational contexts. The task here is to locate which sorts of social-structural, cultural, and social-psychological contexts are more conducive to developing the different modalities of agency that we have outlined in this article. What kinds of settings and situations, for example, tend to keep actors engaged in main-

taining the habitual schematic responses and relations that have become embodied and institutionalized in past experiences? What kinds of contexts provoke or facilitate them toward gaining imaginative distance from those responses and thereby reformulating past patterns through the projection of alternative future trajectories? And finally, what sorts of contexts constrain or enable their capacity for communicative deliberation, by means of which they judge which particular actions are most suitable for resolving the practical dilemmas of emergent situations? The goal here is to locate particular packages of commonly occurring structure-agency relationships, across a wide range of historical, institutional, and interpersonal contexts.²²

We can start in this direction by building upon Swidler's (1986) distinction between "settled" and "unsettled" times. During stable historical periods, she suggests, most people unproblematically employ established cultural competences; however, during periods of upheaval, other forms of agentic activity may come into play. While certain sets of actors might resist change and hold tightly to past routines (such as local or national traditions) in an attempt to ward off uncertainty, others may be more likely to engage in projective activity (as expressed in ideologies and utopias) as they seek to imagine alternative futures for a problematic present. As a countereffect, the strong future orientations provoked by historical change might inhibit actors' responsiveness to situational complexity and practical exigencies (as expressed by ideological rigidity or lack of negotiating capacity). In response, later moments in a historical change cycle might bring more practically evaluative negotiators and institution builders to the fore. We contend that insight into such processes can be gained by looking at the agentic orientations supported by periods of stability and/or change. This recalls what Mead terms the *temporal* dimension of sociality: actors engaged in emergent events find themselves positioned between the old and the new and are thus forced to develop new ways of integrating past and future perspectives. We can formulate this as an exploratory proposition, a probabilistic axis along which to direct empirical research: *Actors who face changing situations that demand (or facilitate) the reconstruction of temporal perspectives can expand their capacity for imaginative and/or deliberative response.*

We can also tackle this question in another way by focusing upon the *relational* (rather than the temporal) dimension of sociality (i.e., the em-

²² We bracket for the purposes of this article the question of how differences in agentic orientation can be empirically measured, although this certainly poses a challenge for future research. We also resist calling agentic orientations "variables" in any linear or causal sense (Abbott 1988), in order to stress the recursivity and multiple determination of all social processes.

beddedness of actors in multiple cultural, social-structural, and social-psychological contexts). A compelling starting place is Rose Laub Coser's (1975, p. 239) elaboration of Merton's theory of the development of individual autonomy from the complexity of role sets: "The multiplicity of expectations faced by the modern individual, incompatible or contradictory as they may be, or rather precisely because they are, makes role articulation possible in a more self-conscious manner than if there were no such multiplicity." The implication here, supported by Coser's research among nursing personnel, is that actors who are located in more complex relational settings must correspondingly learn to take a wider variety of factors into account, to reflect upon alternative paths of action, and to communicate, to negotiate, and to compromise with people of diverse positions and perspectives—all qualities, she argues, that support more autonomous personal and occupational identities (and, by extension, more imaginative and reflective engagements with the contexts of action). Another intriguing research area relates variation in agentic capacity to institutional complexity; for example, in his previously mentioned work on museum reform, DiMaggio (1991) argues that the creation of a professional environment at the *interorganizational* level leads to more critical discourse, formal equality, and purposeful search for alternatives, in contrast to the routine, hierarchy, and scripted forms of rationality that predominate inside organizations. While some researchers have begun to look at how choice making and careers are embedded in complex network interactions (Abbott and Hrycak 1990; Abbott 1997a; Pescosolido 1992), little attention has yet been given to how differently structured networks and careers support variable agentic orientations. We can build upon these findings by formulating another exploratory proposition to serve as a second axis for empirical investigation: *Actors who are positioned at the intersection of multiple temporal-relational contexts can develop greater capacities for creative and critical intervention.*

These formulations may be extended to the study of actors in brokerage positions, long considered an exemplary instance of agentic activity. Such social, political, and economic entrepreneurs seize opportunities for purposive intervention by maneuvering back and forth between different social networks as well as cultural or social-psychological settings. While the critical role of brokers has been well documented by anthropologists, political scientists, economists, and social network analysts (Wolf 1956; Geertz 1960; Boissevain 1974; Marsden 1982; Fernandez and Gould 1994), less attention has been paid to the kinds of temporal constructions (and agentic orientations) that these brokerage positions may support (Mische 1997; Gibson and Mische 1995). Our analysis raises a series of questions in this vein: are actors in such bridging positions more prone to projectivity and practical evaluation than those in more bounded tight-knit con-

texts, given the greater availability of resources for hypothetical rearrangement and comparative evaluation of possible trajectories of action? Does the capacity to draw, when needed, upon different forms of routinized relationships, or conversely, purposively to manipulate, to extend, or to transpose these across changing contexts, underlie their ability to gain greater control and directivity over the various contexts within which they act?

2. *How do changes in agentic orientations allow actors to exercise different forms of mediation over their contexts of action?* This second question requires that we reverse our initial query in order to examine how changes in agentic orientations give actors varying capacities to influence the diverse contexts within which they act. While the foregoing propositions seem to provide a relatively straightforward and optimistic set of scenarios—actors positioned in more temporally and relationally complex settings may have more necessity and/or opportunity to develop the capacity for inventive and deliberative intervention—here we run into greater analytical difficulties. As any student of social processes knows, agentic capacities are only one side of the question; just because actors desire or attempt to intervene does not mean that their interventions will have the desired effects. Both Giddens and Sewell (among others) have taken care to highlight the unintended consequences of action, and a similar point has been made by Marshall Sahlins (1981, 1985, 1991) in a striking series of studies on the interplay between the reproductive and transformative effects of action. While a study of such consequences is by definition exceedingly complex and beyond the scope of this essay, we can alert researchers to some of the paradoxical or counterintuitive situations that a study of agency's interplay with structure might reveal.

For example, an analysis of the multiplex nature of agentic orientations can help to unpack the following paradoxical observation: *Actors who feel creative and deliberative while in the flow of unproblematic trajectories can often be highly reproductive of received contexts.* To understand this phenomenon, we must recall that actors are always simultaneously located in a variety of temporal-relational contexts at once; this is reminiscent of Goffman's (1974) stress upon the multiple embeddings of situations in different frames or vantage points on action. We can extend Goffman's imagery by suggesting that it is possible to be (primarily) iterational in one frame, projective in another, and practical-evaluative in yet a third. Moreover, a switch in frames can reveal apparent contradictions in the reproductive or transformative consequences of action. Take, for example, the case of actors who successfully follow established occupational careers, in which they experience a considerable degree of creative and practical realization. From the perspective of their own professional lives, they are exercising a high degree of personal agency; most likely, they are

highly future oriented in formulating goals and objectives, and well equipped with flexible communicative skills, giving them the capacity to creatively solve emergent problems within the context of the workplace. On the other hand, these same actions can be reframed to show their privileged positioning in relation to other similarly established career trajectories within a particular social-structural matrix at a given historical conjuncture. It may be shown, in fact, that such actors are extremely unquestioning (and iterational) in relation to these larger temporal and relational patterns of action. By "swimming with the current" (Blair-Loy 1997), they unhesitatingly reproduce larger schemas, helping to lock in place social, political, and economic contexts, which, however "unjust" they may appear in an expanded perspective, after all serve the actors well within their own personal and professional lives.

An analysis of shifts in agentic orientations can also shed light upon the converse observation: *Actors who feel blocked in encountering problematic situations can actually be pioneers in exploring and reconstructing contexts of action.* Here again we build from the premise that actors may be capable of switching between agentic orientations and thereby exercising different mediating influences upon their contexts of action. In this case, we can take as an example those who feel that their attempts to follow established trajectories are blocked by the social, political, or economic relations of the day (e.g., the case of women entering male-dominated careers [Blair-Loy 1997], or of members of any excluded group seeking entrance into a previously barred arena). It may be that the reason such border crossers experience difficulties is that they have already projectively expanded and recomposed their proposed fields of action (e.g., the experience of those involved in heady discussions of social reform, such as the civil rights, feminist, or gay and lesbian rights movements), but that when trying to implement those reforms in practice, on either a personal/professional or institutional/legal scale, they encounter hard barriers of interpersonal and institutional conventions. Such actors may not yet have developed the practical-evaluative skills needed to deal with the ambiguities and dilemmas of new and unexpected situations; they may in this case fall back into heavily scripted (or iterational) patterns of interaction, in which conventional roles (e.g., mother, seductress, maiden aunt in the case of women in business careers [Kanter 1977]) are transposed into the new contexts (see also Tilly 1998). On the other hand, as such pioneers make inroads into previously segmented fields, they may also find new and creative ways of fusing, extending, and transforming these received schemas, as they experiment with practical strategies to confront the emergent challenges of historically changing circumstances.

3. *How do actors reconstruct their agentic orientations and thereby alter their own structuring relationships to the contexts of action?* Finally, we

focus upon the research questions opened up by the self-reflexive dimension of agentic orientations, that is, the capacity of actors to reflectively reconstruct their own temporal orientations toward action. In Mead's (1932, p. 72) terms, this is due to the ability of conscious beings to direct attention and intervention toward their own patterns of response: "Life becomes conscious at those points at which the organism's own responses enter into the objective field to which it reacts." Important work in social psychology has focused upon the development of such critical self-awareness, often building upon Meadian conceptions of communicative interaction (Cottrell 1969; Denzin 1988; Callero 1991; Schwalbe 1991). Of particular relevance here is previously mentioned work on life course development, with its focus upon trajectories and turning points (Elder 1985; George 1993), especially work examining the subjective and/or narrative reconstruction of the self through self-interpretive activity during critical life transitions (Cohler 1982). The temporal dimension of such self-construction was stressed three decades ago by Erik Erikson, who showed how conceptions of time develop and change at key transitional periods in the life cycle; for example, a critical task of adolescence is the construction of a sense of a future connected with a past, as manifested in a personal identity that "includes a subjective sense of continuous existence" (Erikson 1968, p. 61). Likewise, researchers on adulthood and aging have noted self-reflective shifts in temporal perspectives as individuals become less preoccupied with the future and more engaged in ruminations upon the past: "While reminiscence is used by much older persons primarily as a means of settling accounts prior to death . . . middle-aged persons are more likely to use reminiscence in an effort to solve problems in the present" (Cohler 1982, p. 225).

With so much attention to temporal perspectives within the subfield of social psychology, it is remarkable that so little of it has made its way into mainstream theoretical and empirical traditions in sociology. More work is necessary in order to link the study of temporal constructions with the varieties of agentic activity that we have tried to delineate in this article. We can formulate this as a final exploratory proposition: *By subjecting their own agentic orientations to imaginative recomposition and critical judgment, actors can loosen themselves from past patterns of interaction and reframe their relationships to existing constraints.* A classic example is Freudian psychoanalysis, in which interpretive recollection of past experiences has a liberating effect upon action; Ricoeur (1970) points out that this process is projective as well, suggesting research into how temporal orientations are intermingled (and undergo changes) in the course of therapeutic processes. Another example is the notion of "cognitive liberation" in the social movement literature (McAdam 1982), in which actors "discover" the possibility of collective action in order to

change an undesired state of affairs. In what ways do such liberating moments require or provoke a recomposition of the temporal construction of the self? Under what conditions do such reconstructions of agentic orientation give actors greater or lesser transformative leverage in relation to their environments? Here we have indicated only a few general components of this process, the full scope and dynamics of which pose ample challenges for future research.

We close with the suggestion that these propositions are not merely relevant for micro- or individual-level analysis but also have important implications for macrolevel research. Abbott (1997b), for example, has suggested that the concept of "turning points" has extensions outside of life course research, including studies of political realignments, business cycles, and scientific progress (not to mention social movements and revolutions).²⁴ We can pose the further query as to whether part of what happens during such periods is a reformulation of the temporal orientations that shape the self-understandings of collective as well as individual actors. Here we echo Aminzade's (1992, p. 470) call for theories "that link the objective temporalities of long-term historical processes to the subjective temporal orientations of social actors." Historical actions and choices are deeply conditioned by how collective actors conceive of the binding power of the past, the malleability of the future, or the capacities of actors to intervene in their immediate situations. Researchers have shown, for example, how cyclical (more iterational) and/or linear (more future-oriented) conceptions of time can place "different limits on the range of adaptive responses to new circumstances" (Aminzade 1992, p. 472; see also Lauer 1973; Goldstone 1987, 1988); such differences in temporal perspectives can have critical effects upon the cohesion or longevity of different forms of community organization and/or collective action (Hall 1978). Yet, despite a few suggestive studies, we still have little understanding of the dynamics by which historical changes in agentic orientations take place. We need further studies of the communicative processes of challenge, experimentation, and debate by which actors formulate new temporally constructed understandings of their own abilities to engage in individual and collective change, as well as how these microlevel processes intersect with longer-term social, political, and economic trajectories.

²⁴ Abbott, however, is less interested in the subjective composition of such turning points than in the structural characteristics that make them particularly susceptible to transformative action. Trajectories, he claims, can be conceived of as narratively constructed "networks through time," linked by occasional transitions that bring about a reformulation of the logic governing the connection between past and future possibilities. Turning points are the "peculiarly essential junctures . . . where action might make particularly consequential bridges by making or breaking links between many networks" (Abbott 1997b, p. 99).

Such an approach would have the additional merit of placing the discussion of agency squarely within the context of its own essential historicity.

CONCLUSION

We have argued throughout this essay that human agency needs to be radically reconceptualized. Neither rational choice theory, norm-based approaches, nor any of the other sociological perspectives extant today provide a fully adequate understanding of its significance and constituent features. Nor do such perspectives satisfactorily answer the question as to how agency interpenetrates with and impacts upon the temporal-relational contexts of action.

We have contended that one key to understanding the variable orientations of agency toward its structural contexts lies in a more adequate theorization of the temporal nature of human experience. Actors are always living simultaneously in the past, future, and present, and adjusting the various temporalities of their empirical existence to one another (and to their empirical circumstances) in more or less imaginative or reflective ways. They continuously engage patterns and repertoires from the past, project hypothetical pathways forward in time, and adjust their actions to the exigencies of emerging situations. Moreover, there are times and places when actors are more oriented toward the past, more directive toward the future, or more evaluative of the present; actors may switch between (and reflexively transform) their orientations toward action, thereby changing their degrees of flexible, inventive, and critical response toward structuring contexts. Such a perspective lays the basis for a richer and more dynamic understanding of the capacity that actors have to mediate the structuring contexts within which action unfolds. We have referred to this perspective as *relational pragmatics*.

Finally, this point of view also opens up the possibility to conceive of moral and practical issues regarding human freedom, creativity, and democracy in a more satisfactory and powerful way. In this essay, we have not laid out a normative theory that actually distinguishes between "better" or "worse" agentic processes, "more or less morally worthy" projects. The elaboration of such a theory would require even longer and more complex arguments than those presented here. Yet, we have delineated the analytical space within which reflective and morally responsible action might be said to unfold. Throughout, we have stressed the reconstructive, (self-) transformative potentialities of human agency, when faced with contradictory or otherwise problematic situations. What are commonly referred to as norms and values, we can now add, are themselves by-products of actors' engagement with one another in ambiguous and challenging circumstances; they emerge when individuals experience a dis-

cordance between the claims of multiple normative commitments. Problematic situations of a moral and practical nature can thus become resolved (to the extent that they can become resolved at all [Hook 1974]) only when actors reconstruct the temporal-relational contexts within which they are embedded and, in the process, transform their own values and themselves. As Mead (1964, p. 149) expresses it, "The appearance of . . . different interests in the forum of reflection [leads to] the reconstruction of the social world, and the consequent appearance of the new self that answers to the new object."

While the optimistic progressivism of the classical pragmatists may appear relatively simple and even naive from our position at the close of the 20th century, the orientation toward action that they present still resonates powerfully as we attempt to respond to a rapidly changing world composed of increasingly complex and overlapping matrices of social, political, and economic relations. If we cannot control the consequences of our interventions, we can at least commit ourselves to a responsive, experimental, and deliberative attitude as we confront emergent problems and possibilities across the variety of contexts within which we act. As the pragmatist thinkers never tired of reminding us, this is a preeminently dialogic and communicative process, which unfolds in perpetual interaction with the social universe. Both the pragmatist conception of the responsive intelligence and the Kantian ideal of the enlarged mentality can be of use to us in the continuing challenge to develop ever more comprehensive, cosmopolitan, and universalistic perspectives—perspectives nevertheless flexible enough to respond to situational complexity and ambiguity. The "mode of associated living" that Joas (1996), following Dewey, calls "creative democracy" embodies such moral intelligence on a transpersonal scale; it involves "conjoint communicated experience" (Dewey 1980, p. 93) in which imaginative reformulation and practical reasoning are undertaken in common through inquiry into moral and practical problems on the model of an experimental science. If our perspective on human agency does not in itself resolve such problems, it can at least help to give social science a more adequate theoretical grounding, so that it can become a creative and vital participant in this democratic debate.

REFERENCES

- Abbott, Andrew 1988 "Transcending General Linear Reality." *Sociological Theory* 6:169–86
- . 1997a. "Of Time and Space: The Contemporary Relevance of the Chicago School." *Social Forces* 75 1149–82.
- . 1997b. "On the Concept of Turning Point." *Comparative Social Research* 16 85–105.
- Abbott, Andrew, and Alexandra Hrycak. 1990. "Measuring Resemblance in Sequence

- Data: An Optimal Matching Analysis of Musicians' Careers." *American Journal of Sociology* 96:144-85.
- Alexander, Jeffrey C. 1988a. *Action and Its Environments*. New York: Columbia University Press.
- , ed. 1988b. *Durkheimian Sociology: Cultural Studies*. New York: Cambridge University Press.
- . 1992. "Some Remarks on 'Agency' in Recent Sociological Theory." *Perspectives* (Theory Section Newsletter, American Sociological Association) 15:1-4.
- . 1995. "The Reality of Reduction: The Failed Synthesis of Pierre Bourdieu." Pp. 128-217 in *Fin de Siècle Social Theory: Relativism, Reduction, and the Problem of Reason*. London: Verso.
- Aminzade, Ronald. 1992. "Historical Sociology and Time." *Sociological Method and Research* 20:456-80.
- Aquinas, Thomas. 1948. *Summa Theologica*, 3 vols. Translated by fathers of the English Dominican Province. Westminster, Md.: Christian Classics.
- Archer, Margaret S. 1982. "Morphogenesis versus Structuration: On Combining Structure and Action." *British Journal of Sociology* 33:455-83.
- . 1988. *Culture and Agency: The Place of Culture in Social Theory*. Cambridge: Cambridge University Press.
- Arendt, Hannah. 1977a. "The Crisis in Culture: Its Social and Political Significance." Pp. 197-226 in *Between Past and Future*. Harmondsworth: Penguin Books.
- . 1977b. *The Life of the Mind*. New York: Harcourt, Brace Jovanovich.
- . 1984. "Thinking and Moral Considerations: A Lecture." *Social Research* 51:7-37.
- Aristotle. 1985. *Nicomachean Ethics*, translated by Terence Irwin. Indianapolis, Ind.: Hackett.
- Axelrod, Robert. 1984. *The Evolution of Cooperation*. New York: Basic.
- Bakhtin, M. M. 1986. "The Problem of Speech Genres." Pp. 60-102 in *Speech Genres and Other Late Essays*, edited by Caryl Emerson and Michael Holquist. Translated by Vern W. McGee. Austin: University of Texas Press.
- Benhabib, Seyla. 1987. "The Generalized Other and the Concrete Other: The Kohlberg-Gilligan Controversy and Feminist Theory." Pp. 77-95 in *Feminism as Critique*, edited by Seyla Benhabib and Drucilla Cornell. Minneapolis: University of Minnesota Press.
- . 1992a. "In the Shadow of Aristotle and Hegel: Communicative Ethics and Current Controversies in Practical Philosophy." Pp. 23-67 in *Situating the Self*. New York: Routledge.
- . 1992b. "Judgment and the Moral Foundations of Politics in Hannah Arendt's Thought." Pp. 121-44 in *Situating the Self*. New York: Routledge.
- . 1992c. *Situating the Self*. New York: Routledge.
- Berger, Peter, and Thomas Luckmann. 1966. *The Social Construction of Reality: A Treatise in the Sociology of Knowledge*. New York: Doubleday.
- Bergson, H. 1989. *Time and Free Will*, translated by F. L. Pogson. London: Allen Unwin.
- Bernstein, Richard J. 1971. *Praxis and Action*. Philadelphia: University of Pennsylvania Press.
- Berteaux, Daniel, ed. 1981. *Biography and History: The Life History Approach in the Human Sciences*. Beverly Hills, Calif.: Sage.
- Blair-Loy, Mary. 1997. "Career Patterns of Executive Women in Finance. An Optimal Matching Analysis." Manuscript. Washington State University, Department of Sociology.
- Boussevain, Jeremy. 1974. *Friends of Friends: Networks, Manipulators, and Coalitions*. Oxford: Blackwell.

- Bourdieu, Pierre. 1977. *Outline of a Theory of Practice*, translated by Richard Nice. Cambridge: Cambridge University Press.
- . 1984. *Distinction: A Social Critique of the Judgment of Taste*, translated by Richard Nice. Cambridge, Mass.: Harvard University Press.
- . 1990. *The Logic of Practice*, translated by Richard Nice. Stanford, Calif.: Stanford University Press.
- Bourdieu, Pierre, and Loïc Wacquant. 1992. *An Invitation to Reflexive Sociology*. Chicago: University of Chicago Press.
- Brint, Steven, and Jerome Karabel. 1991. "Institutional Origins and Transformations: The Case of American Community Colleges." Pp. 337–60 in *The New Institutionalism in Organizational Analysis*, edited by Walter Powell and Paul J. DiMaggio. Chicago: University of Chicago Press.
- Bruner, Jerome. 1986. *Actual Minds, Possible Worlds*. Cambridge, Mass.: Harvard University Press.
- Callero, Peter L. 1991. "Toward a Sociology of Cognition." Pp. 43–54 in *The Self-Society Dynamic: Cognition, Emotion, and Action*, edited by Judith A. Howard and Peter L. Callero. Cambridge: Cambridge University Press.
- Camic, Charles. 1986. "The Matter of Habit." *American Journal of Sociology* 91:1039–87.
- Carr, David. 1986. *Time, Narrative, and History*. Bloomington: Indiana University Press.
- Castoriadis, Cornelius. 1987. *The Imaginary Constitution of Society*, translated by Kathleen Blamey. Cambridge, Mass.: MIT Press.
- Cohen, Jean L. 1985. "Strategy or Identity: New Theoretical Paradigms and Contemporary Social Movements." *Social Research* 52:663–716.
- Cohen, Jean L., and Andrew Arato. 1992. *Civil Society and Political Theory*. Cambridge, Mass.: MIT Press.
- Cohler, Bertram J. 1982. "Personal Narrative and Life Course." *Life Course Development and Behavior* 4:206–41.
- Cohn, Norman. 1977. *The Pursuit of the Millennium: Revolutionary Millenarians and Mystical Anarchists of the Middle Ages*. New York: Oxford University Press.
- Colapietro, Vincent M. 1990. "The Vanishing Subject of Contemporary Discourse: A Pragmatic Response." *Journal of Philosophy* 87:644–55.
- Coleman, James S. 1986. "Social Theory, Social Research, and a Theory of Action." *American Journal of Sociology* 6:1309–35.
- . 1990. *Foundations of Social Theory*. Cambridge, Mass.: Harvard University Press.
- Cook, Karen Sweeers, and Margaret Levi, eds. 1990. *The Limits of Rationality*. Chicago: University of Chicago Press.
- Corsaro, William A. 1992. "Interpretive Reproduction in Children's Peer Culture." *Social Psychology Quarterly* 55:160–77.
- Coser, Rose Laub. 1975. "The Complexity of Roles as a Seedbed of Individual Autonomy." Pp. 237–63 in *The Idea of Social Structure. Essays in Honor of Robert Merton*, edited by Lewis Coser. New York: Harcourt Brace Jovanovich.
- Cottle, T. J., and Klineberg, S. 1974. *The Present of Things Future*. New York: Free Press.
- Cottrell, L. 1969. "Interpersonal Interaction and the Development of the Self." Pp. 543–70 in *Handbook of Socialization Theory and Research*, edited by D. Goslin. Chicago: Rand McNally.
- Davies, Brian. 1992. *The Thought of Thomas Aquinas*. Oxford: Clarendon Press.
- De Certeau, Michel. 1984. *The Practice of Everyday Life*, translated by Steven Rendall. Berkeley and Los Angeles: University of California Press.
- Denzin, Norman K. 1985. "Act, Language and Self in Symbolic Interactionist Thought." *Studies in Symbolic Interactionism* 9:51–80.

- Dearoche, Henri. 1979. *The Sociology of Hope*, translated by Carol Martin-Sperry. London: Routledge & Kegan Paul.
- Devolder, M. L., and W. Lens. 1982. "Academic Achievement and Future Time Perspective: A Cognitive-Motivational Concept." *Journal of Personality and Social Psychology* 42:566-71.
- Dewey, John. 1922. *Human Nature and Conduct*. New York: Modern Library.
- . 1940. "Creative Democracy: The Task before Us." Pp. 220-28 in *The Philosopher of the Common Man. Essays in Honor of John Dewey to Celebrate His Eightieth Birthday*, edited by Sidney Ratner. New York: Greenwood.
- . 1978. *Ethics*, edited by Jo Ann Boydston. Carbondale: Southern Illinois University Press.
- . 1980. *Democracy and Education*. Vol. 9 of *The Middle Works of John Dewey, 1899-1924*. Edited by Jo Ann Boydston. Carbondale: Southern Illinois University Press.
- . 1981. "The Need for a Recovery of Philosophy." Pp. 58-97 in *The Philosophy of John Dewey*, edited by John J. McDermott. Chicago: University of Chicago Press.
- . 1985. "The Logic of Judgments of Practice." Pp. 14-82 in *John Dewey: Essays on Education and Politics, 1915*, edited by Jo Ann Boydston. Carbondale: Southern Illinois University Press.
- DiMaggio, Paul J. 1988. "Interest and Agency in Institutional Theory." Pp. 3-21 in *Institutional Patterns and Organizations: Culture and Environment*, edited by Lynne G. Zucker. Cambridge, Mass.: Ballinger.
- . 1991. "Constructing an Organizational Field as a Professional Project: U.S. Art Museums, 1920-1940." Pp. 267-92 in *The New Institutionalism in Organizational Analysis*, edited by Walter W. Powell and Paul J. DiMaggio. Chicago: University of Chicago Press.
- Douglas, Mary. 1985. *Purity and Danger*. London: Ark Paperbacks.
- Durkheim, Émile. 1961. *Moral Education. A Study in the Theory and Application of the Sociology of Education*, translated by Everett K. Wilson and Herman Schnurer. New York: Free Press.
- Elder, Glenn H., Jr. 1974. *Children of the Great Depression: Social Change in Life Experiences*. Chicago: University of Chicago Press.
- . 1985. "Perspectives on the Life Course." Pp. 23-49 in *Life Course Dynamics*, edited by Glenn H. Elder, Jr. Ithaca, N.Y.: Cornell University Press.
- . 1994. "Time, Human Agency, and Social Change: Perspectives on the Life Course." *Social Psychological Quarterly* 57:4-15.
- Elster, Jon. 1989. *The Cement of Society: A Study of Social Order*. Cambridge: Cambridge University Press.
- Emirbayer, Mustafa. 1992a. "Beyond Structuralism and Voluntarism: The Politics and Discourse of Progressive School Reform, 1890-1930." *Theory and Society* 21: 621-64.
- . 1992b. "The Shaping of a Virtuous Citizenry: Educational Reform in Massachusetts, 1830-1860." *Studies in American Political Development* 6:391-419.
- . 1997. "Manifesto for a Relational Sociology." *American Journal of Sociology* 103:281-317.
- Emirbayer, Mustafa, and Jeff Goodwin. 1994. "Network Analysis, Culture, and the Problem of Agency." *American Journal of Sociology* 99:1411-53.
- . 1996. "Symbols, Positions, and Objects: Toward a New Theory of Revolution and Collective Action." *History and Theory* 35:358-74.
- Emirbayer, Mustafa, and Mimi Sheller. 1996. "Publics in History." Paper presented at the meeting of the Social Science History Association, New Orleans.
- Erikson, Erik H. 1968. *Identity: Youth and Crisis*. New York: Norton.
- Fernandez, Roberto M., and Roger V. Gould. 1994. "A Dilemma of State Power: Bro-

- kerage and Influence in the Mental Health Policy Domain." *American Journal of Sociology* 99:1455-91.
- Fligstein, Neil. 1991. "The Structural Transformation of American Industry: An Institutional Account of the Causes of Diversification in the Largest Firms, 1919-1979." Pp. 311-36 in *The New Institutionalism in Organizational Analysis*, edited by Walter W. Powell and Paul J. DiMaggio. Chicago: University of Chicago Press.
- Galaskiewicz, Joseph. 1991. "Making Corporate Actors Accountable: Institution-Building in Minneapolis-St. Paul." Pp. 293-310 in *The New Institutionalism in Organizational Analysis*, edited by Walter W. Powell and Paul J. DiMaggio. Chicago: University of Chicago Press.
- Gamson, William A. 1992. *Talking Politics*. Cambridge: Cambridge University Press.
- Garfinkel, Harold. 1963. "A Conception of and Experiments with 'Trust' as a Condition of Concerted Stable Actions." Pp. 187-238 in *Motivation and Social Interaction*, edited by O. J. Harvey. New York: Ronald Press.
- . 1984. *Studies in Ethnomethodology*. Cambridge: Polity Press.
- Garton Ash, Timothy. 1990. *The Magic Lantern: The Revolution of '89 Witnessed in Warsaw, Budapest, Berlin, and Prague*. New York: Random House.
- Geertz, Clifford. 1960. "The Javanese Kijaji. The Changing Role of a Cultural Broker." *Comparative Studies in Society and History* 2:228-49.
- George, Linda K. 1993. "Sociological Perspectives on Life Transitions." *Annual Review of Sociology* 19:353-73.
- Gergen, Kenneth J. 1994. *Realities and Relationships: Soundings in Social Construction*. Cambridge, Mass.: Harvard University Press.
- Gergen, Kenneth J., and Mary M. Gergen. 1983. "Narratives of the Self." Pp. 254-73 in *Studies in Social Identity*, edited by Theodore R. Sarbin and Karl Scheibe. New York: Praeger.
- . 1984. "The Social Construction of Narrative Accounts." Pp. 173-89 in *Historical Social Psychology*, edited by Kenneth J. Gergen and Mary M. Gergen. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- . 1988. "Narrative and the Self as Relationship." *Advances in Experimental Social Psychology* 21:17-56.
- Gibson, David, and Ann Mische. 1995. "Internetwork Encounters and the Emergence of Leadership." Pre-Print Series, no. 202. Lazarsfeld Center for the Social Sciences, Columbia University.
- Giddens, Anthony. 1976. *New Rules for Sociological Method*. New York: Basic.
- . 1979. *Central Problems in Social Theory*. Berkeley and Los Angeles: University of California Press.
- . 1984. *The Constitution of Society*. Berkeley and Los Angeles: University of California Press.
- . 1991. *Modernity and Self-Identity: Self and Society in the Late Modern Age*. Stanford, Calif.: Stanford University Press.
- Gilligan, Carol. 1982. *In a Different Voice: Psychological Theory and Women's Development*. Cambridge, Mass.: Harvard University Press.
- Goffman, Erving. 1959. *The Presentation of Self in Everyday Life*. New York: Doubleday.
- . 1974. *Frame Analysis: An Essay on the Organization of Experience*. New York: Harper & Row.
- Goldstone, Jack A. 1987. "Cultural Orthodoxy, Risk, and Innovation: The Divergence of East and West in the Early Modern World." *Sociological Theory* 5:119-35.
- . 1988. "East and West in the Seventeenth Century: Political Crises in Stuart England, Ottoman Turkey, and Ming China." *Comparative Studies in Society and History* 30:103-42.
- Goodwyn, Lawrence. 1991. *Breaking the Barrier: The Rise of Solidarity in Poland*. New York: Oxford University Press.

- Greene, A. L. 1986. "Future Time Perspective in Adolescence: The Present of Things Future, Revisited." *Journal of Youth and Adolescence* 15:99-113.
- . 1990. "Great Expectations: Constructions of the Life Course during Adolescence." *Journal of Youth and Adolescence* 19:289-306.
- Habermas, Jürgen. 1984-89. *The Theory of Communicative Action*, 2 vols. Translated by Thomas McCarthy. Boston: Beacon Press.
- . 1990. *Moral Consciousness and Communicative Action*, translated by Christian Lenhardt and Sherry Weber Nicholsen. Cambridge, Mass.: MIT Press.
- . 1993. *Justification and Application*, translated by Clara P. Cronin. Cambridge, Mass.: MIT Press.
- Halbwachs, Maurice. 1992. *On Collective Memory*, edited and translated by Lewis Coser. Chicago: University of Chicago Press.
- Hall, John R. 1978. *The Ways Out: Utopian Communal Groups in an Age of Babylon*. Boston: Routledge & Kegan Paul.
- Haraway, Donna. 1988. "Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective." *Feminist Studies* 14:575-99.
- Hays, Sharon. 1994. "Structure and Agency and the Sticky Problem of Culture." *Sociological Theory* 12:57-72.
- Hechter, Michael. 1992. "Should Values Be Written Out of the Sociologist's Lexicon?" *Sociological Theory* 10:214-30.
- . 1994. "The Role of Values in Rational Choice Theory." *Rationality and Society* 6:318-33.
- Heidegger, Martin. 1962. *Being and Time*, translated by John Macquarrie and Edward Robinson. New York: Harper & Row.
- Hobsbawm, Erik J., and Terence Ranger, eds. 1983. *The Invention of Tradition*. New York: Cambridge University Press.
- Homans, George Caspar. 1964. "Bringing Men Back In." *American Sociological Review* 29:809-18.
- Hook, Sydney. 1974. "Pragmatism and the Tragic Sense of Life." Pp. 3-25 in *Pragmatism and the Tragic Sense of Life*. New York: Basic.
- Husserl, Edmund. 1960. *Cartesian Meditations: An Introduction to Phenomenology*, translated by Dorion Cairns. The Hague: Martinus Nijhoff.
- James, William. 1976. *Essays in Radical Empiricism*. Cambridge, Mass.: Harvard University Press.
- Joas, Hans. 1985. *G. H. Mead: A Contemporary Re-Examination of His Thought*. Cambridge, Mass.: MIT Press.
- . 1993. *Pragmatism and Social Theory*. Chicago: University of Chicago Press.
- . 1996. *The Creativity of Action*, translated by Jeremy Gaines and Paul Keast. Chicago: University of Chicago Press.
- . N.d. "The Creativity of Action: Pragmatism and the Critique of the Rational Actor Model." Manuscript. Free University, Berlin, John F. Kennedy Institute for North American Study.
- Kant, Immanuel. 1951. *Critique of Judgment*, translated by J. H. Bernard. New York: Hafner.
- . 1956. *Critique of Practical Reason*, translated by Lewis White Beck. Indianapolis, Ind.: Bobbs-Merrill.
- . 1964. *Groundwork of the Metaphysics of Morals*, translated by H. J. Paton. New York: Harper Torchbooks.
- . 1965. *Critique of Pure Reason*, translated by Norman Kemp Smith. New York: St. Martin's.
- . 1971. *The Metaphysics of Morals*. Part 2, *Doctrine of Virtue*, translated by Mary J. Gregor. Philadelphia: University of Pennsylvania Press.
- Kanter, Rosabeth. 1977. *Men and Women of the Corporation*. New York: Basic.

- Kearney, Richard. 1988. *The Wake of Imagination: Toward a Postmodern Culture*. Minneapolis: University of Minnesota Press.
- Kestenbaum, Victor. 1977. *The Phenomenological Sense of John Dewey: Habit and Meaning*. Atlantic Highlands, N.J.: Humanities Press.
- Kierkegaard, Søren. 1944. *Concluding Unscientific Postscript*, translated by David F. Swenson. Completed after his death by Walter Lowrie. Princeton, N.J.: Princeton University Press.
- Kohlberg, Lawrence. 1981. *The Philosophy of Moral Development*. Vol. 1, *Essays on Moral Development*. New York: Harper & Row.
- Kohli, Martin. 1986. "Social Organization and Subjective Construction of the Life Course." Pp. 271–92 in *Human Development and the Life Course: Multidisciplinary Perspectives*, edited by Aage B. Sørensen, Franz E. Weinert, Lonnie R. Sherrod. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Laba, Roman. 1991. *The Roots of Solidarity: A Political Sociology of Poland's Working-Class Democratization*. Princeton, N.J.: Princeton University Press.
- Lacan, Jacques. 1977. *Écrits: A Selection*, translated by Alan Sheridan. New York: Norton.
- Lamont, Michèle. 1992. *Money, Morals, and Manners*. Chicago: University of Chicago Press.
- Lauer, Robert H. 1973. "Temporality and Social Change: The Case of 19th Century China and Japan." *Sociological Quarterly* 14:451–64.
- Leifer, Eric M. 1988. "Interaction Preludes to Role Setting." *American Sociological Review* 53:865–78.
- Levine, Donald N. 1985. *The Flight from Ambiguity: Essays in Social and Cultural Theory*. Chicago: University of Chicago Press.
- Lévi-Strauss, Claude. 1966. *The Savage Mind*. Chicago: University of Chicago Press.
- Lewin, Kurt. 1948. "Time Perspective and Morale." Pp. 48–70 in *Civilian Morale*, edited by Goodwin Watson. New York: Reynal, Hitchcock.
- Lipset, Seymour Martin, Martin A. Trow, and James S. Coleman. 1962. *Union Democracy: The Internal Politics of the International Typographical Union*. Garden City, N.Y.: Anchor Books.
- Locke, John. 1978. *Two Treatises of Government*. New York: E. P. Dutton.
- Luhmann, Niklas. 1990. *Essays on Self-Reference*. New York: Columbia University Press.
- Lukes, Steven. 1973. *Individualism*. Oxford: Basil Blackwell.
- MacIntyre, Alasdair. 1981. *After Virtue: A Study in Moral Theory*. Notre Dame, Ind.: University of Notre Dame Press.
- Mansbridge, Jane J. 1983. *Beyond Adversary Democracy*. Chicago: University of Chicago Press.
- March, James G. 1978. "Bounded Rationality, Ambiguity, and the Engineering of Choice." *Bell Journal of Economics* 9:587–608.
- March, James G., and Johan P. Olsen. 1976. *Ambiguity and Choice in Organizations*. Bergen: Universitetsforlaget.
- . 1984. "The New Institutionalism: Organizational Factors in Political Life." *American Political Science Review* 78:734–49.
- March, James G., and Herbert A. Simon. 1958. *Organizations*. New York: Wiley.
- Marsden, Peter V. 1982. "Brokerage Behavior in Restricted Exchange Networks." Pp. 201–18 in *Social Structure and Network Analysis*, edited by Peter V. Marsden and Nan Lin. Beverly Hills, Calif.: Sage.
- Marx, Karl. 1978. *The Marx-Engels Reader*, 2d ed. Edited by Robert C. Tucker. New York: Norton.
- McAdam, Doug. 1982. *Political Process and the Development of Black Insurgency: 1930–1970*. Chicago: University of Chicago Press.

- McLoughlin, William G. 1978. *Revivals, Awakenings, and Reform: An Essay on Religion and Social Change in America, 1607-1977*. Chicago: University of Chicago Press.
- Mead, George Herbert. 1932. *The Philosophy of the Present*. Chicago: University of Chicago Press.
- . 1934. *Mind, Self, and Society: From the Standpoint of a Social Behaviorist*. Chicago: University of Chicago Press.
- . 1938. *The Philosophy of the Act*. Chicago: University of Chicago Press.
- . 1964. "The Social Self." Pp. 142-49 in *George Herbert Mead: Selected Writings*, edited by Andrew J. Reck. Indianapolis, Ind.: Bobbs-Merrill.
- Melucci, Alberto. 1989. *Nomads of the Present: Social Movements and Individual Needs in Contemporary Society*, edited by John Keane and Paul Mier. Philadelphia: Temple University Press.
- . 1994. "The Process of Collective Identity." Pp. 41-63 in *Social Movements and Culture*, edited by Hank Johnston and Bert Klandermans. Minneapolis: University of Minnesota Press.
- Merleau-Ponty, Maurice. 1964. *Sense and Non-Sense*, translated by Hubert L. Dreyfus and Patricia Allen Dreyfus. Evanston, Ill.: Northwestern University Press.
- Meyer, John W. 1986. "The Self and the Life Course." Pp. 199-217 in *Human Development and the Life Course: Multidisciplinary Perspectives*, edited by Aage B. Sørensen, Franz E. Weinert, and Lonnie R. Sherrod. Hillsdale, N.J.: Lawrence Erlbaum.
- Meyer, John W., John Boli, and George M. Thomas. 1987. "Ontology and Rationalization in the Western Cultural Account." Pp. 12-37 in *Institutional Structure: Constituting State, Society, and the Individual*, edited by George M. Thomas, John W. Meyer, and John Boli. Beverly Hills, Calif.: Sage.
- Meyer, John W., and Brian Rowan. 1991. "Institutionalized Organizations: Formal Structure as Myth and Ceremony." Pp. 41-62 in *The New Institutionalism in Organizational Analysis*, edited by Walter W. Powell and Paul J. DiMaggio. Chicago: University of Chicago Press.
- Meyer, John W., W. R. Scott, and T. E. Deal. 1981. "Institutional and Technical Sources of Organizational Structure." Pp. 151-78 in *Organization and the Human Services*, edited by H. D. Stein. Philadelphia: Temple University Press.
- Mische, Ann. 1997. "Projects, Identities, and Social Networks: Brazilian Youth Mobilization and the Making of Civic Culture." Paper presented at the annual meeting of the American Sociological Association, Toronto.
- Moore, Barrington, Jr. 1978. *Injustice: The Social Bases of Obedience and Revolt*. Armonk, N.Y.: M. E. Sharpe.
- Morris, Aldon D., and Carol McClurg Mueller, eds. 1992. *Frontiers in Social Movement Theory*. New Haven, Conn.: Yale University Press.
- Münch, Richard. 1981. "Talcott Parsons and the Theory of Action. I. The Structure of the Kantian Core." *American Journal of Sociology* 86:709-39.
- . 1994. *Sociological Theory*, 3 vols. Chicago: Nelson-Hall.
- Nussbaum, Martha C. 1986. *The Fragility of Goodness*. Cambridge: Cambridge University Press.
- Nystrom, N. C., and William H. Starbuck. 1984. "To Avoid Organizational Crisis, Unlearn." *Organizational Dynamics* 13:53-65.
- O'Donnell, Guillermo, and Philippe Schmitter. 1986. *Transitions from Authoritarian Rule*. Vol. 4, *Tentative Conclusions about Uncertain Democracies*. Baltimore: Johns Hopkins University Press.
- Olick, Jeffrey K. 1997. "The Textual Dynamics of Memory: Continuities and Departures in German Commemoration of May 8th, 1945." Manuscript. Columbia University, Department of Sociology.
- Olick, Jeffrey K., and Daniel Levy. 1997. "Collective Memory and Cultural Con-

- straint: Holocaust Myth and Rationality in German Politics." *American Sociological Review* 62:921-36.
- O'Rand, Angela M., and Margaret L. Kreckler. 1990. "Concepts of the Life Cycle: Their History, Meanings, and Uses in the Social Sciences." *Annual Review of Sociology* 16:241-62.
- Ortner, Sherry B. 1984. "Theory in Anthropology since the Sixties." *Comparative Studies in Society and History* 26 (1): 126-66.
- Padgett, John F., and Christopher K. Ansell. 1993. "Robust Action and the Rise of the Medici, 1400-1434." *American Journal of Sociology* 98:1259-1319.
- Parsons, Talcott. 1968. *The Structure of Social Action*, 2 vols. New York: Free Press.
- Parsons, Talcott, and Edward Shils. 1951. *Toward a General Theory of Action*. Cambridge, Mass.: Harvard University Press.
- Patterson, Orlando. 1991. *Freedom*. Vol. 1, *Freedom in the Making of Western Culture*. New York: Basic.
- Pescosolido, Bernice A. 1992. "Beyond Rational Choice. The Social Dynamics of How People Seek Help." *American Journal of Sociology* 97:1096-1138.
- Pizzorno, Alessandro. 1986. "Some Other Kind of Otherness." Pp. 355-73 in *Development, Democracy, and the Art of Trespassing: Essays in Honor of Albert O. Hirschman*, edited by Alejandro Foxley, Michael S. McPherson, and Guillermo O'Donnell. Notre Dame, Ind.: University of Notre Dame Press.
- Powell, Walter W. 1986. "How the Past Informs the Present: The Uses and Liabilities of Organizational Memory." Paper presented at the Conference on Communication and Collective Memory. University of Southern California, Annenberg School.
- Powell, Walter W., and Paul J. DiMaggio, eds. 1991. *The New Institutionalism in Organizational Analysis*. Chicago: University of Chicago Press.
- Ricoeur, Paul. 1970. *Freud and Philosophy: An Essay on Interpretation*, translated by Denis Savage. New Haven, Conn.: Yale University Press.
- . 1991. *From Text to Action*, translated by Kathleen Blamey and John B. Thompson. Evanston, Ill.: Northwestern University Press.
- Riegel, Klaus F. 1975. "Adult-Life Crises: Toward a Dialectic Interpretation of Development." Pp. 99-128 in *Life-Span Developmental Psychology: Normative Life Crisis*, edited by Nancy Datan and L. H. Ginsberg. New York: Academic Press.
- . 1977. "Dialectics of Time." Pp. 3-45 in *Life-Span Developmental Psychology: Dialectical Perspectives on Experimental Research*, edited by Nancy Datan and Hayne W. Reese. New York: Academic Press.
- Roth, Guenther, and Wolfgang Schluchter. 1979. *Max Weber's Vision of History: Ethics and Methods*. Berkeley and Los Angeles: University of California Press.
- Sacks, Harvey, Emmanuel A. Schegloff, and Gail Jefferson. 1974. "A Simplest Systematics for the Organization of Turn-Taking for Conversation." *Language* 50:696-735.
- Sahlins, Marshall. 1981. *Historical Metaphors and Practical Reason*. Ann Arbor: University of Michigan Press.
- . 1985. *Islands of History*. Chicago: University of Chicago Press.
- . 1991. "The Return of the Event, Again: With Reflections on the Beginnings of the Great Fijian War of 1944 to 1955 between the Kingdoms of Bau and Rewa." In *Clio in Oceania*, edited by A. Biersack. Washington, D.C.: Smithsonian Institution Press.
- Sarbin, Theodore R., ed. 1986. *Narrative Psychology: The Storied Nature of Human Conduct*. New York: Praeger.
- Sartre, Jean-Paul. 1956. *Being and Nothingness: An Essay in Phenomenological Ontology*, translated by Hazel E. Barnes. Secaucus, N.J.: Citadel.
- Schutz, Alfred. 1962. "Choosing among Projects of Action." Pp. 67-96 in *Collected Papers*. Vol. 1, *The Problem of Social Reality*, edited by Maurice Natanson. The Hague: Martinus Nijhoff.

- . 1964. "Tiresias, or Our Knowledge of Future Events." Pp. 277–93 in *Collected Papers*. Vol. 2, *Studies in Social Theory*, edited by Arvid Brodersen. The Hague: Martinus Nijhoff.
- . 1967. *The Phenomenology of the Social World*, translated by George Walsh and Frederick Lehnert. Evanston, Ill.: Northwestern University Press.
- . 1978. *The Theory of Social Action: The Correspondence of Alfred Schutz and Talcott Parsons*, edited by Richard Grathoff. Bloomington: Indiana University Press.
- Schwalbe, Michael L. 1991. "Social Structure and the Moral Self." Pp. 281–304 in *The Self-Society Dynamic: Cognition, Emotion, and Action*, edited by Judith A. Howard and Peter L. Callero. Cambridge: Cambridge University Press.
- Schwartz, Barry. 1991. "Social Change and Collective Memory: The Democratization of George Washington." *American Sociological Review* 56:221–36.
- Scott, James C. 1985. *Weapons of the Weak: Everyday Forms of Peasant Resistance*. New Haven, Conn.: Yale University Press.
- . 1990. *Domination and the Arts of Resistance: Hidden Transcripts*. New Haven, Conn.: Yale University Press.
- Sewell, William H., Jr. 1992. "A Theory of Structure: Duality, Agency, and Transformation." *American Journal of Sociology* 98:1–29.
- Snow, David A., and Robert D. Benford. 1988. "Ideology, Frame Resonance, and Participant Mobilization." *International Social Movement Research* 1:197–217.
- . 1992. "Master Frames and Cycles of Protest." Pp. 133–55 in *Frontiers in Social Movement Theory*, edited by Aldon D. Morris and Carol McClurg Mueller. New Haven, Conn.: Yale University Press.
- Snow, David A., E. Burke Rochford, Jr., Steven K. Worden, and Robert D. Benford. 1986. "Frame Alignment Processes, Micromobilization, and Movement Participation." *American Sociological Review* 51:464–81.
- Somers, Margaret. 1992. "Narrativity, Narrative Identity, and Social Action: Rethinking English Working-Class Formation." *Social Science History* 16 (4): 591–630.
- Sorokin, Pitirim. 1947. *Society, Culture, and Personality: Their Structure and Dynamics, a System of General Sociology*. New York: Harper.
- Steinberg, Marc W. 1996. "Toward a Dialogic Analysis of Social Movement Culture: Reframing Frame Analysis." Paper presented at the meeting of the Social Science History Association, New Orleans.
- Stepan, Alfred. 1978. "Political Leadership and Regime Breakdown: Brazil." Pp. 110–37 in *The Breakdown of Democratic Regimes: Latin America*, edited by Juan J. Linz and Alfred Stepan. Baltimore: Johns Hopkins University Press.
- Swidler, Ann. 1986. "Culture in Action. Symbols and Strategies." *American Sociological Review* 51:273–86.
- Tarrow, Sidney. 1992. "Mentalities, Political Cultures, and Collective Action Frames: Constructing Meanings through Action." Pp. 174–202 in *Frontiers in Social Movement Theory*, edited by Aldon D. Morris and Carol McClurg Mueller. New Haven, Conn.: Yale University Press.
- Taylor, Charles. 1985. "What Is Human Agency?" Pp. 15–44 in *Philosophical Papers*. Vol. 1, *Human Agency and Language*. Cambridge: Cambridge University Press.
- . 1989. *Sources of the Self*. Cambridge, Mass.: Harvard University Press.
- . 1991. "The Dialogical Self." Pp. 304–14 in *The Interpretive Turn: Philosophy, Science, Culture*, edited by David R. Hilley, James F. Bohman, and Richard Shusterman. Ithaca, N.Y.: Cornell University Press.
- . 1993. "To Follow a Rule . . ." Pp. 45–60 in *Bourdieu: Critical Perspectives*, edited by Craig Calhoun, Edward LiPuma, and Moishe Postone. Chicago: University of Chicago Press.
- Thomas, W. I., and Florian Znaniecki. 1918. *The Polish Peasant in Europe and America*. New York: Knopf.

- Thompson, Edward Palmer. 1966. *The Making of the English Working Class*. New York: Vintage.
- . 1993. *The Witness against the Beast: William Blake and the Moral Law*. New York: New Press.
- Tilly, Charles. 1986. *The Contentious French*. Cambridge, Mass.: Harvard University Press.
- . 1994. "History and Sociological Imagining." *Tocqueville Review* 15:57–74.
- . 1998. *Durable Inequality*. Berkeley and Los Angeles: University of California Press, in press.
- Touraine, Alain, Francois Dubet, Michel Wlewiorka, and Jan Strzelecki. 1983. *Solidarity: Poland 1980–81*, translated by David Denby. Cambridge: Cambridge University Press.
- Trotsky, Leon. 1980. *The History of the Russian Revolution*. New York: Pathfinder.
- Turner, Stephen. 1994. *The Social Theory of Practices: Tradition, Tacit Knowledge, and Presuppositions*. Chicago: University of Chicago Press.
- Turner, Victor. 1974. *Dramas, Fields, and Metaphors: Symbolic Action in Human Society*. Ithaca, N.Y.: Cornell University Press.
- Wacquant, Loïc. 1992a. "Toward a Social Praxeology. The Structure and Logic of Bourdieu's Sociology." Pp. 1–59 in *An Invitation to Reflexive Sociology*, by Pierre Bourdieu and Loïc Wacquant. Chicago: University of Chicago Press.
- . 1992b. "The Social Logic of Boxing in Black Chicago." *Sociology of Sports Journal* 9:221–54.
- Walzer, Michael. 1965. *The Revolution of the Saints: A Study in the Origins of Radical Politics*. Cambridge, Mass.: Harvard University Press.
- . 1980. "A Theory of Revolutions." Pp. 201–23 in *Radical Principles: Reflections of an Unreconstructed Democrat*. New York: Basic.
- . 1985. *Exodus and Revolution*. New York: Basic.
- Weber, Max. 1946. "Politics as a Vocation." Pp. 77–128 in *From Max Weber*, edited by Hans Gerth and C. Wright Mills. New York: Oxford University Press.
- . 1978. *Economy and Society*, 2 vols. Edited by Guenther Roth and Claus Wittich. Berkeley: University of California Press.
- Wessman, Alden E., and Bernard S. Gorman. 1977. "The Emergence of Human Awareness and Concepts of Time." Pp. 4–55 in *The Personal Experience of Time*, edited by Bernard S. Gorman and Alden E. Wessman. New York: Plenum.
- Westbrook, Robert B. 1991. *John Dewey and American Democracy*. Ithaca, N.Y.: Cornell University Press.
- White, Harrison. 1992. *Identity and Control*. Princeton, N.J.: Princeton University Press.
- . 1995. "Passages reticulaires, acteurs, et grammaire de la domination." *Revue française de sociologie* 36:705–23.
- Wiley, Norbert. 1994. *The Semiotic Self*. Chicago: University of Chicago Press.
- Willis, Paul. 1977. *Learning to Labor: How Working Class Kids Get Working Class Jobs*. New York: Columbia University Press.
- Wolf, Eric R. 1956. "Aspects of Groups Relations in a Complex Society: Mexico." *American Anthropologist* 58:1065–78.
- Zucker, Lynne G. 1977. "The Role of Institutionalization in Cultural Persistence." *American Sociological Review* 42:726–43.
- . 1983. "Organizations as Institutions." Pp. 1–47 in *Research in the Sociology of Organizations*, vol. 2. Edited by Samuel B. Bacharach. Greenwich, Conn.: JAI Press.

The Long-Term Effects of Foreign Investment Dependence on Economic Growth, 1940–1990¹

Jeffrey Kentor
Johns Hopkins University

Does dependence of a national economy on foreign investment promote economic growth or underdevelopment? The recent exchanges between Glenn Firebaugh and William J. Dixon and Terry Boswell suggest the contentious nature of this debate. This study analyzes models similar to those tested in previous research but with data from earlier time points to examine the long-term effects of foreign capital penetration. Accumulated stocks of foreign capital/GDP in 1938 have a short-term (five-year) positive effect on economic growth followed by a 20-year lagged negative effect on economic growth beginning in 1960 and lasting at least 30 years. Similar effects are found using a second indicator of foreign investment dependence, debits on investment income, for the 1950–90 period.

The international migration of capital has facilitated development of the world's natural resources and has been instrumental in transmitting the direct effects of the industrial revolution from area to area. Thereby it has helped to increase the quantities and varieties of goods and services generally available and has raised living standards for some or most of the world's populations. (Cleona Lewis, *Debtor and Creditor Countries*, 1948)

Following an initial growth spurt this (foreign investment) will create an industrial structure in which monopoly is predominant, labor is insufficiently absorbed, and there is underutilization of the productive forces. Thus, the peripheral countries that adopt this path of uneven development based on income inequality and foreign capital imports will experience economic stagnation, under- and unemployment, and increasing marginalization of the population relative to countries that are less

¹I would like to thank Christopher Chase-Dunn for his comments and support. Thanks also to Volker Bornschler for his comments on an earlier draft and to Peter Grimes for the use of his data set. This work was significantly improved by the comments of the *AJS* reviewers. Direct all correspondence to Jeffrey Kentor, Department of Sociology, Johns Hopkins University, Baltimore, Maryland 21218.

penetrated by transnational corporations and relative to the growth potential of the country. (Volker Bornschier and Christopher Chase-Dunn, *Transnational Corporations and Underdevelopment*, 1985)

The above quotes summarize the debate that has continued for more than 20 years. Does dependence of a national economy on foreign investment promote economic growth or underdevelopment? The answer to this question has far-reaching implications. A country's level of economic development has an obvious impact on nearly all aspects of social life, so it is worthwhile to examine the mechanisms that have contributed to the historically large and growing economic inequalities existing across the core-periphery distribution of countries (Korzeniewicz and Moran 1997).

The classical modernization perspective, as reflected in Lewis's (1948) comments above, argued that the export of capital to undeveloped countries promoted economic growth by creating industries, transferring technology, and fostering a "modern" perspective in the local population.

Dependency theory (Amin 1974; Frank 1979) challenged this pervasive belief. The dependency school argued that ownership of capital determined its effect on the underdeveloped economy. An economy controlled by foreign interests would not develop organically. It would grow in a disarticulated manner. The natural linkages that would evolve from locally controlled capital would not occur. Profits would be exported. The interests of the ruling elite would be allied with those of owners of the foreign capital. Income inequality would grow. The economy would stagnate.

The empirical work of Chase-Dunn (1975) and Bornschier and Chase-Dunn (1985) supported the above theoretical argument. They found that their measure of foreign capital penetration (PEN: a ratio of foreign investments to total capital stocks) in 1967 had a negative effect on GNP per capita growth, 1965–75. Bornschier and Chase-Dunn characterized these results within Wallerstein's (1974, 1979, 1980, 1989) world-systems perspective, arguing that foreign capital structured peripheral economies in the interests of core economies and in a manner that reproduced the core-periphery division of labor.

Subsequently, a diverse literature exploring the negative effects of foreign capital emerged along a spectrum of economic, social, and political dimensions including income inequality (Chase Dunn 1975a; Bornschier 1980), unemployment (Bornschier and Ballmer-Cao 1979), overurbanization (Kentor 1981; Timberlake and Kentor 1983), fertility rates (London 1988), rebellions (Boswell and Dixon 1990), among many others. There

appeared to be clear agreement that penetration of foreign capital promoted a host of negative effects on the penetrated economy.

The validity of Bornschier and Chase-Dunn's findings was questioned by Firebaugh (1992), who argued that these negative effects were spurious, a statistical artifact of the "denominator effect" of PEN. Firebaugh concluded from his reanalysis of Bornschier and Chase-Dunn's data that penetration of foreign capital has a positive effect on economic growth but one that is smaller than the positive effect that domestic capital investment has.

Dixon and Boswell (1996) addressed Firebaugh's concerns by constructing two new measures of foreign capital penetration: the ratios of (1) foreign stocks to total capital stocks and (2) foreign stocks to gross domestic product (GDP). Their reanalyses of the data with these new measures of foreign investment dependence supported Bornschier and Chase-Dunn's earlier findings of the negative effect of foreign investment dependence.

Dixon and Boswell (1996) pursued their critique of Firebaugh (1992) on a theoretical level. They argued that Firebaugh failed to distinguish between foreign capital investment and foreign investment dependence. It is not foreign capital per se that has a negative effect on economic growth but rather the control of foreign capital over the host economy that has a negative impact. Dixon and Boswell outlined the mechanisms by which penetration of foreign capital retards economic growth. They distinguished between the "differential productivity" associated with foreign investment—the disarticulated growth that occurs as the peripheral economy is reoriented toward the interests of the core—and the "negative externalities" of foreign capital penetration—the social and political consequences of this restructuring that goes beyond the economic "drag" of foreign investment, such as overurbanization and income inequality. Firebaugh (1996) responded to Dixon and Boswell's critique by asserting that they failed to account for the ratio problem, arguing that the substitution of GDP for total capital stocks in the denominator of their measure of foreign capital penetration was not meaningful since the correlation of GDP and total capital stocks is .99 for the countries studied.

A serious problem with both of these studies, however, is that they examine the effects of foreign capital penetration over a relatively short time period, 1967–73. The mechanisms by which these negative effects are hypothesized to occur would seem to suggest longer time frames. Bornschier and Chase-Dunn (1985) hypothesized that foreign investment dependence would have a short-term positive effect on the host economy as capital flows generated employment and purchases that represented the start-up costs of the new ventures. This would be followed by a longer term nega-

tive effect as the host economy is restructured for the benefit of the core. This distinction between short- and long-term effects anticipates Dixon and Boswell's theoretical understanding of the mechanisms by which foreign capital exerts negative effects on penetrated economies.

Fortunately, earlier data are available from two sources to examine these long-term effects: the Brookings Institution, which compiled cross-national data on accumulated foreign stocks in 1938, and the International Monetary Fund (IMF), which provides data on debits on investment income for 1938 and then annually beginning in 1946. I will use these two alternative measures to examine the long-term effects of foreign capital penetration on economic growth in peripheral countries.

THE ANALYSES

I examine the effects of dependence on foreign capital on economic growth over a longer time period than previously used, 1938–90, to separate the differential short-term and long-term effects hypothesized by Bornschier and Chase-Dunn (1985). Alternative measures of foreign capital penetration and GDP are used, along with different case bases and two methods of handling missing data, to provide independent confirmation of the results.

Measures

The dependent variable of economic growth is measured as the natural log of GDP per capita (GDP/pc). These data are taken from Grimes (1996) and Maddison (1995). The use of GDP data in historical cross-national research is somewhat problematic. Methods of calculating GDP vary by country and time period. Exchange rates fluctuate. The GDP data may not accurately reflect relative standards of living due to the differential purchasing power across countries, especially in the periphery of the world-economy. For this reason, I use two different measures of GDP. Grimes (1996) gives GDP data in constant 1990 dollars, while Maddison (1995) provides purchasing power parity (PPP) GDP estimates.² The analyses are performed with Grimes's GDP data since he gives data on more peripheral countries in 1940. The analyses are then rerun using Maddison's data.³

²The correlations between Grimes's (1996) and Maddison's (1995) GDP measures are .85 in 1950, .97 in 1970, and .96 in 1990.

³The data for 1940 were estimated for four countries by applying Maddison's (1995) GDP indices to Grimes's (1996) 1945 GDP data.

The first measure of foreign investment dependence, *accumulated foreign stocks 1938*, is taken from Lewis (1948). These data are somewhat incomplete. For some countries only total stock values are given, while for others the data are broken down as portfolio or direct investments as well as whether corporate or government based. It would be preferable to examine these different forms of foreign investment separately since it is argued that direct investments give foreign interests the greatest control and, hence, should have the strongest negative effect. In this case, the data were only complete enough to subtract government loans from total stocks. This measure is standardized by GDP, which is consistent with Dixon and Boswell (1996) and logged to normalize the skewed distribution. This measure will be referred to as "LFANGS38GDP" (\log_e [foreign accumulated nongovernment stocks]).

The second measure of foreign investment dependence is *debits on investment income* (LDEB49GDP), taken from the IMF *Balance of Payments Yearbook* (1949, 1951) and first used as an indicator of foreign capital penetration by Chase-Dunn (1975*b*). These data represent income accruing to foreigners on enterprises controlled by them in the host country, including net profits transferred abroad or reinvested. Undistributed profits, interest, and dividends accrued but not transferred are also included. These figures include both direct and portfolio investments. Debits on investment income should reflect the level of foreign investment, given a relatively constant profit rate. I have averaged these data over the three-year period 1947–49 to take yearly fluctuations into account. For six countries, I used 1950 data, which was the earliest data available. Inclusion of these cases is justified, given the stable nature of foreign investment. These data were also standardized by GDP and logged to normalize the skewed distribution.

An examination of the correlations among the various measures of foreign investment provides an interesting starting point for the empirical analyses. The correlation between Lewis's 1938 data on accumulated foreign stocks and the IMF's data on debits on investment income in 1938 is .88, while the correlation between debits on investment income in 1938 and 1949 is .75. Of greatest interest, however, is the correlation between debits on investment income 1949 and accumulated foreign stocks in 1967 (Bornschiefer and Heintz 1979), which is also .75. These figures are remarkable, given the different aspects of foreign investment being measured, the inherent differences in data collection and a 30-year span. This reflects a significant stability in foreign investment, both across measures and over time. The correlation matrix for these variables is given in the appendix. Exports/GDP and total GDP are included in these analyses as control variables. Inclusion of these control variables is consistent with

prior research (Dixon and Boswell 1996; Bornschier and Chase-Dunn 1985).⁴

Countries Included in the Analyses

As dependency theory focuses on the effects of foreign investment dependence on peripheral countries, it would be preferable to exclude core countries from the analyses. Unfortunately, this was not possible with the LFANGS38GDP data as there are only 36 countries with complete data on all variables, approximately half of which are peripheral countries. Also included in these analyses are five nonmarket countries.

The situation for LDEB49GDP is better. High GDP/pc and nonmarket countries are excluded from the data base, leaving 75 possible cases in the analyses. A listwise deletion of missing data leaves us with 33 cases to be analyzed. These results will be compared with results obtained with analyses employing a maximum-likelihood estimation (MLE) of missing data on the entire 75 country data set. A list of countries included in the analyses is given in appendix table A2.

Research Design

In the first part of the analyses, the effects of the two measures of foreign capital penetration were estimated with two separate series of panel regression models in which the dependent variable at T_2 was regressed on the lagged dependent variable at T_1 , along with the other independent variables. So there are 10 panel regressions for the LFANGS38GDP measure, beginning with 1940–45, 1940–50, 1940–55, and so on, ending with the 1940–90 time period, and eight regressions for the LDEB49GDP measure. The resulting coefficients of the panel regressions indicate the effects of foreign investment on per capita GDP growth over the given time period.⁵

⁴To replicate earlier models more closely, it would have been useful to include a control measure of domestic capital formation. Unfortunately these data are not available for 1940. The inclusion of GDP in the regression equations controls for total capital stocks since these two variables are so highly correlated, .99 according to Firebaugh (1996). For a substantive discussion of why these control variables are included, see Bornschier and Chase-Dunn (1985, p. 92).

⁵I have chosen to employ a panel regression design instead of the growth scores used in prior research. Inclusion of the lagged dependent variable (GDP/pc) controls for any nonlinear specification in this variable prior to that time. These analyses were also run using a change score of GDP per capita growth as the dependent variable. These findings were consistent with the panel regression model reported below.

In the second part of the analyses, a structural equation model of these effects is estimated on the complete database of 75 peripheral countries with a MLE of missing data. The rationale and methodology will be discussed in some detail below. I also include the control variables of exports/GDP and GDP in 1950 and 1965 in this analysis to better specify the model.

RESULTS

Panel Regression Analyses

The results for the first series of panel regression examining the effects of LFANGS38GDP on economic growth are presented in table 1 and offer clear support for the earlier findings of the negative effects of foreign capital penetration on economic growth.⁶

My measure of foreign capital penetration in 1938 (LFANGS38GDP) has an initial short-term positive effect on GDP per capita growth, 1940–45 ($\beta = .10$). Between 1960 and 1990, however, the results indicate a continuing significant lagged negative effect of foreign capital penetration on economic growth. This negative effect becomes significant in the 1940–60 period ($\beta = -.05$) and increases through the 1940–90 period ($\beta = -.18$).

A second series of panel regressions, not shown here, were run on the above model excluding high GDP/pc countries, leaving 20 cases in the analysis. The results of these analyses were consistent with those above, although the effects were larger. The results of the next series of panel regressions, using LDEB49GDP, are given in table 2. The findings are comparable to those using LFANGS38GDP, except that the initial positive effect of LDEB49pc on economic development is not significant. There is a 20-year lagged negative effect of LDEB49pc on economic growth beginning in 1970 and persisting through 1990 ($\beta = -.31$).⁷

A scatterplot of the relationship between GDP/pc growth 1950–75 and foreign investment dependence/GDP 1949 is presented in figure 1 and provides a dramatic illustration of this relationship. Japan has the lowest level of foreign investment dependence/GDP and the highest GDP/pc growth rate during this period. Other countries exhibiting low levels of foreign investment dependence/GDP and high economic growth include Greece, Taiwan, and Israel. Those countries exhibiting high levels of foreign investment dependence/GDP and relatively low economic growth include Honduras, Chile, Sri Lanka, and Egypt among others.

⁶ The inclusion of the lagged dependent variable in panel regression models can result in artificially high R^2 and betas greater than 1.0.

⁷ No significant differences were found when these analyses were rerun using Maddison's (1995) measure of GDP.

TABLE 1
OLS ESTIMATES OF EFFECTS OF LEFANGS38GDP ON ECONOMIC GROWTH (GDP/PC), 1940-90

	1945			1950			1955			1960			1965		
	b	β	t-ratio	b	β	t-ratio	b	β	t-ratio	b	β	t-ratio	b	β	t-ratio
LGDPpc40	1.04	1.04	17.4***	1.1	1.06	29***	1.06	1.04	25.6***	1.01	.99	23.3***	.99	.96	17.92***
LEFANGS38GDP	1	15	2.58***	.02	.03	.76	-.02	-.03	-.69	-.07	-.1	-2.42*	-.11	-.15	-2.92***
EXPM40GDP	1.3E-05	.08	1.6	1.9E-05	12	3.7***	2.5E-05	16	4.54***	3.8E-05	2	5.12***	3.4E-05	22	4.52***
LGDP1940	.05	.07	1.18	-.01	-.01	-.2	7.3E-05	1.1E-05	0	.01	.01	.24	.01	.02	.36
Constant	-1.64		-2.36*	-.86		-1.96	-.13		-.27	.89		1.78	1.51		2.35*
Adjusted R	.95		.98			.97			.97			.96			.96
<hr/>															
	1970			1975			1980			1985			1990		
	b	β	t-ratio	b	β	t-ratio	b	β	t-ratio	b	β	t-ratio	b	β	t-ratio
LGDPpc40	.94	.91	14.5***	.89	.88	11.26***	.85	.88	10.16***	.85	.82	8.84***	.93	.67	6.11***
LEFANGS38GDP	-.15	-.211	-3.47***	-.18	-.26	-3.38***	-.18	-.27	-3.16***	-.22	-.31	-3.43***	-.22	-.23	-2.17*
EXPM40GDP	3.8E-05	.24	4.33***	4.2E-05	.27	3.93***	4.2E-05	.28	3.6***	4.8E-05	.3	3.68***	-4E-05	-.19	-1.98*
LGDP1940	.02	.03	.49	.02	.03	.32	-.01	-.01	-.09	.02	.03	.31	-.04	-.04	-.35
Constant	2.41		3.18***	3.19		3.46***	3.59		3.57***	4.09		3.67***	.89		1.78
Adjusted R	.94		.89			.88			.85			.85	.97		.97

NOTE—N = 32

* P < .05

*** P < .01.

TABLE 2
OLS ESTIMATES OF EFFECTS OF LDEB49GDP ON ECONOMIC GROWTH (GDP/PC), 1950-90

	1955			1960			1965			1970		
	b	β	t-ratio	b	β	t-ratio	b	β	t-ratio	b	β	t-ratio
LGDPpc50	1.02	.98	49.14**	1.01	.97	33**	1.04	.95	22.46**	1.08	.93	19**
LDEB49GDP . . .	-.002	-.004	-2	-.02	-.04	-1.13	-.05	-.08	-1.8	-.09	-.14	-2.68**
EXF50GDP	-.0004	-.02	-.92	-.0003	-.01	-.482	-.0004	-.02	-4	.0004	.02	.36
LGDP195001	.01	.63	.03	.04	1.22	.02	.03	.58	.05	.06	1.19
Constant02		.17	.14		.65	.13		.41	-.07		-.17
Adjusted R99			.98			.95			.94		
	1975			1980			1985			1990		
	b	β	t-ratio	b	β	t-ratio	b	β	t-ratio	b	β	t-ratio
LGDPpc40	1.09	.91	16.19**	1.08	.9	14.71**	1.07	.87	12**	1.09	.83	10.49**
LFANGS38GDP	-.1	-.16	-2.6**	-.12	-.18	-2.77**	-.16	-.23	-2.93**	-.2	-.28	-3.21**
EXP40GDP	-.0004	-.001	-.03	.0005	.02	.35	.001	.06	.72	.002	.08	.98
LGDP194003	.05	.77	.04	.06	.88	.07	.09	1.16	.1	.12	1.39
Constant06	.14	.89	.21		.4	.19		.31	-.4.97		-6.91**
Adjusted R92			.9			.87			.84		

NOTE.—N = 35

* P < .05

** P < .01

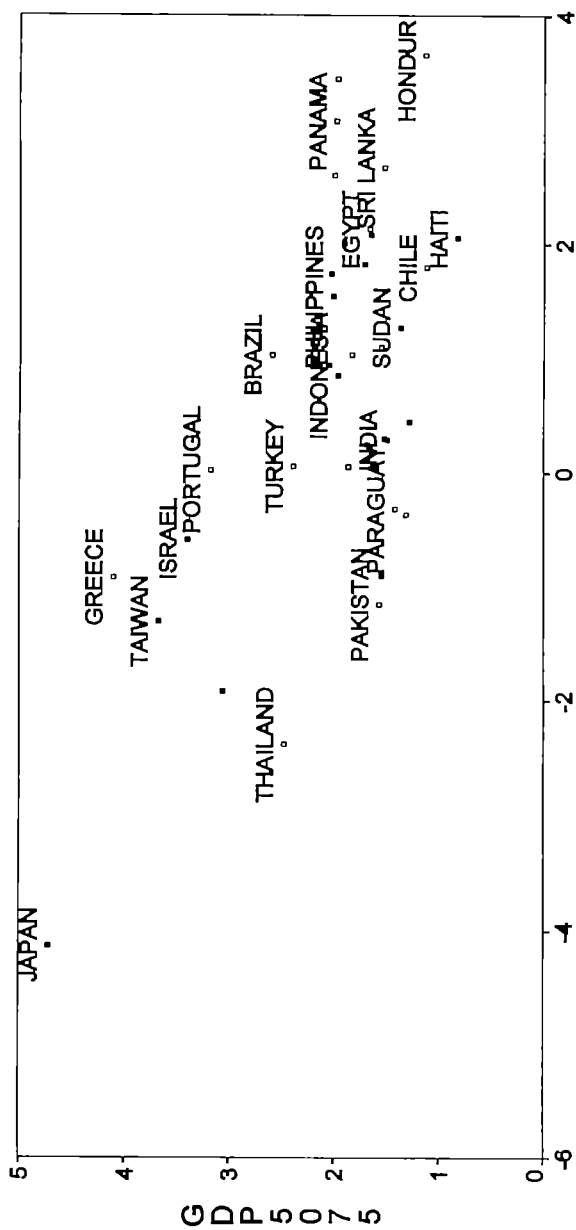


Fig. 1.—Scatterplot of GDP/pc growth, 1950–79, by debits on investment income/GDP 1949

The scatterplot of the relationship between GDP/pc growth 1950–90 and foreign investment dependence, shown in figure 2, reveals movement in the distribution over this 40-year period. Although Japan still exhibits an extremely high growth rate, the relative growth rates of Taiwan, Greece, and Israel have decreased while Thailand's relative growth rate has increased. An initial question that arises from examination of these scatterplots is whether the extreme scores for Japan have a significant impact on the regression analyses reported above. They do not. Exclusion of Japan from the analyses does not have a significant effect on the findings.

While the inclusion of Japan may be problematic on a statistical level, its inclusion is important on a theoretical basis since it provides the clearest example of the relationship between foreign investment and economic growth. The relatively low level of foreign investment is used by Arrighi (1994) to explain Japan's spectacular rise in the world-economy. Japan's low level of foreign investment and associated rapid economic growth were primarily functions of external controls, according to Arrighi, a process referred to as "sponsored mobility." Arrighi points out that the United States actually protected the Japanese market from transnational corporation penetration, thereby allowing locally controlled industries to develop. This was done in the national security interests of the United States, which desired an economically strong Japanese state to act as a buffer against China.

A comparison of the effects of LFANGS38GDP and LDEB49GDP on economic growth by lag periods is given in table 3. The standardized coefficients are nearly equivalent, the only difference being the absence of a significant five-year positive effect of LDEB49GDP on GDP/pc. The 20-year to 40-year lagged negative effects are remarkably similar between the two measures of foreign investment dependence.

Structural Equation Model

In this section, a structural equation model is estimated to provide additional support for the above analyses. The primary concern addressed here deals with the significant amount of missing data in the data set. It also provides a convenient method of including control variables at later time points in the analyses.

The issue of missing data merits some discussion, as it is a significant concern with this type of research.⁸ In this data set, for example, fewer

⁸ For a useful summary of current methods of handling missing data, see Little and Rubin (1989).

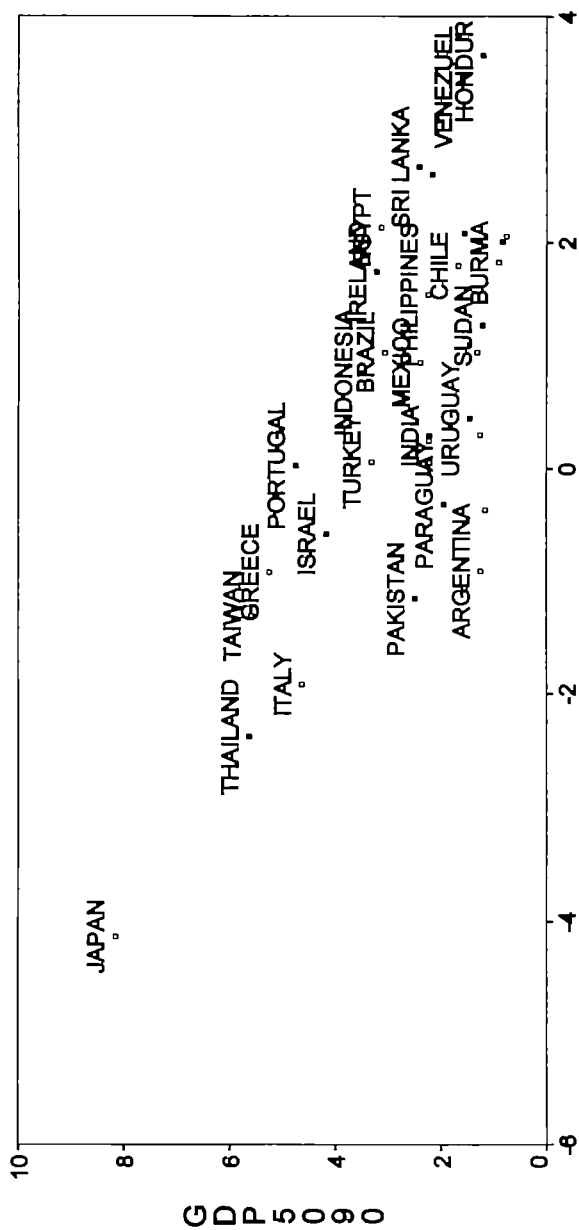


FIG. 2.—Scatterplot of GDP/pc growth, 1950–90, by debits on investment income/GDP 1949

TABLE 3
COMPARISON OF THE EFFECTS OF LFANGS38GDP AND LDEB49GDP ON ECONOMIC GROWTH, 1940-90, STANDARDIZED
REGRESSION COEFFICIENTS

	5-Year Lag	10-Year Lag	15-Year Lag	20-Year Lag	25-Year Lag	30-Year Lag	35-Year Lag	40-Year Lag	45-Year Lag	50-Year Lag
LFANGS38GDP15**	.03	-.03	-.1*	-.15**	-.21**	-.26**	-.27**	-.31**	-.23*
LDEB49GDP	0	-.03	-.08	-.14**	-.16**	-.18**	-.23**	-.28**		

* $P < .05$
** $P < .01$

than half the 75 peripheral countries have complete data for all variables. These missing data may not be randomly distributed. Peripheral countries are more likely to have missing data, and these are also the countries more likely to be affected by foreign investment dependence. The two most common methods of handling missing data are listwise and pairwise deletion (Kim and Curry 1977). Neither method is without problems. Listwise deletion discards available data, while pairwise deletion may not represent the underlying population especially if the requirement of random missing data cannot be met.

Arbuckle (1996) compares these methods with MLEs of missing data.⁹ He demonstrates that MLEs are more consistent and less biased than listwise or pairwise deletion, especially when the data do not meet the requirement of missing completely at random or even missing at random. The structural equation modeling program used in these analyses (AMOS) uses this maximum-likelihood method of estimating missing data, providing a useful comparison with the results obtained in the previous section using listwise deletion.

The model to be estimated is presented in figure 3. In this model, debits on investment income have direct effects on GDP/pc in 1960, 1975, and 1990. Exports/GDP 1950 and GDP 1950 have direct effects on GDP/pc in 1960. The only substantive addition to the prior model here is the inclusion of exports/GDP and GDP/pc in 1960, 1975, and 1990. These variables have direct effects on GDP/pc in 1975 and 1990. Inclusion of these variables controls for relevant changes over the time period in question. The results for the theoretically relevant variables in the above model are presented in figure 4 and the MLEs for the entire model are given in table 4. LDEB49GDP has a significant lagged negative effect on economic growth that increases over time with standardized coefficients of $-.01$ in 1960 (not significant), $-.16$ in 1975, and $-.15$ in 1990. These effects are net of the effects of GDP and exports/GDP in 1950 and 1965.

DISCUSSION

These results provide clear support for the prior findings of Bornschier and Chase-Dunn (1985) and Dixon and Boswell (1996). The short-term positive effect of foreign investment on economic growth, generated by

⁹ The maximum-likelihood approach of dealing with missing data differs from listwise or pairwise deletion of data in that information is not excluded in the estimation process. While there are alternative methods of calculating MLEs, the conceptual process is the same. First, a model is specified. Next, the probability density (the *likelihood*) of the observed data for this model is calculated. Finally, parameters are estimated to maximize this likelihood.

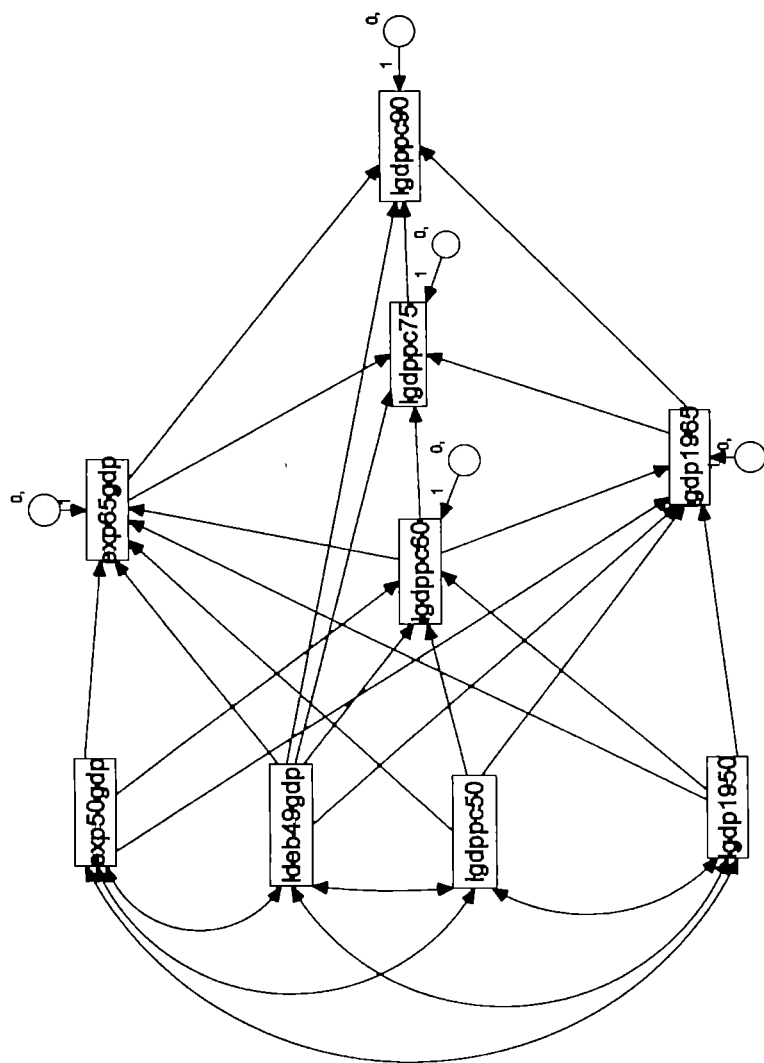


FIG. 3.—A structural equations model of the effects of debits on investment income/GDP in 1949 on economic growth between 1950 and 1990.

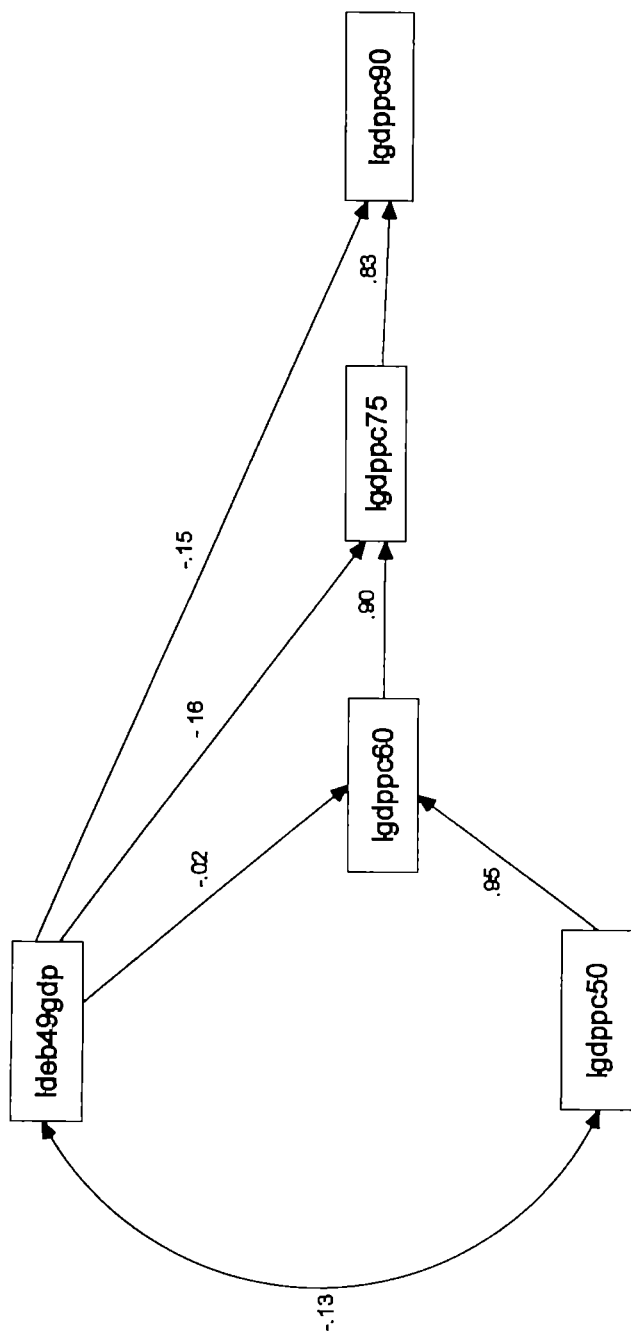


FIG. 4.—MLEs (standardized) of the effects of debits on investment income/GDP in 1949 on economic growth between 1950 and 1990, controlling for the effects of GDP and exports/GDP in 1950 and 1965 (not shown).

TABLE 4

MLEs OF THE EFFECTS OF DEBITS ON INVESTMENT INCOME 1949 ON ECONOMIC GROWTH, 1950-90

	Estimate	SE	Critical Ratio	Standardized Coefficient
Regression weights				
LGDPpc60 \leftarrow LDEB49GDP	-.013	.024	-.532	-.025
LGDPpc60 \leftarrow EXP50GDP	0	.001	-.157	-.007
LGDPpc60 \leftarrow LGDPpc50	1.013	.038	29.96	.946
LGDPpc60 \leftarrow LGDP1950022	.023	.966	.038
LGDP1965 \leftarrow LGDP1950973	.018	54.855	.934
LGDP1965 \leftarrow EXP50GDP001	.001	.789	.015
LGDP1965 \leftarrow LGDPpc50	-.828	.101	-8.204	-.43
LGDP1965 \leftarrow LDEB49GDP	-.03	.018	-1.676	-.033
LGDP1965 \leftarrow LGDPpc60901	.095	9.439	.5
EXP65GDP \leftarrow LGDP1950	-4.71	.954	-4.939	-.577
EXP65GDP \leftarrow EXP50GDP071	.035	2.043	.253
EXP65GDP \leftarrow LGDPpc50	-6.632	5.573	-1.19	-.439
EXP65GDP \leftarrow LDEB49GDP502	.952	.527	.07
EXP65GDP \leftarrow LGDPpc60	13.567	5.261	2.579	.962
LGDPpc75 \leftarrow LGDPpc60	1.076	.037	29.229	.903
LGDPpc75 \leftarrow LDEB49GDP	-.096	.018	-5.414	-.158
LGDPpc75 \leftarrow EXP65GDP01	.003	3.606	.12
LGDPpc75 \leftarrow LGDP1965023	.024	.977	.035
LGDPpc90 \leftarrow LDEB49GDP	-.101	.033	-3.012	-.15
LGDPpc90 \leftarrow EXP65GDP01	.005	1.825	.104
LGDPpc90 \leftarrow LGDP1965071	.041	1.719	.098
LGDPpc90 \leftarrow LGDPpc75918	.058	15.919	.832
Covariance				
LDEB49g \longleftrightarrow LGDPpc50	-.214	.214	-1	
EXP50GDP \longleftrightarrow LGDP1950	-24.504	10.145	-2.415	
LGDPpc50 \longleftrightarrow LGDP1950577	.171	3.385	
LDEB49g \longleftrightarrow LGDP1950	-.938	.399	-2.349	
LDEB49g \longleftrightarrow EXP50GDP	37.298	13.018	2.865	
EXP50GDP \longleftrightarrow LGDPpc50	-13.505	5.615	-2.405	
Intercepts:				
LGDPpc6014	.262	.532	
LGDP1965041	.205	.199	
EXP65GDP	-26.854	10.898	-2.464	
LGDPpc75	-.285	.211	-1.348	
LGDPpc90524	.339	1.545	

capital inflows and increased employment, is replaced by a consistent long-term lagged negative effect beginning 20 years later and lasting at least 30 years. The strongest support is found in table 2 of the analyses. Two different measures of foreign investment dependence, measured at different times and analyzed over different cases, provide nearly equivalent results.

The only anomaly between the LFANGS38GDP and LDEB49GDP analyses was the absence of an initial positive effect of LDEB49GDP on GDP/pc in 1955. The most likely reason for this is a reflection of the characteristics of this measure. New enterprises do not usually generate profits immediately. Rather, there is a lag time of anywhere from one to four years before new ventures become profitable. LDEB49GDP is an average of profits from 1947 to 1949. As such, it may actually reflect accumulated stocks as early as 1942. When GDP/pc 1940 is used as the lagged dependent variable with GDP/pc 1950 as the dependent variable, the coefficient is positive and 1.5 times the size of its standard error. This suggests that foreign investment does, in fact, have an initial positive effect on economic growth.

One of the most interesting findings of this research is that the negative effects of foreign investment dependence do not appear to diminish over time. How are we to interpret this? I would argue that these persistent negative effects reflect two aspects of foreign investment dependence. First, it is a reflection of the stability of foreign investment over time. The infrastructure that develops in a country dominated by foreign capital is conducive to subsequent investment. The social and political structure that evolves, a small local elite whose interests are linked with those of the foreign investors, would ease the entry of future foreign investment. The economic infrastructure created to deal with issues of foreign property ownership, currency convertibility, and labor laws in a manner favorable to foreign interests would also make subsequent investment more likely.

There is empirical support for this argument, as indicated by the high correlation among the various measures of investment dependence used here with accumulated foreign stocks in 1967, as discussed earlier. A panel regression estimating the effects of LDEB49GDP on accumulated foreign stocks in 1967 (AFS67GDP), presented in table 5, provides additional support. LDEB49GDP has a significant positive effect AFS67GDP, with an unstandardized coefficient of .37.

Further, this apparent continuous negative effect of foreign capital penetration on economic growth reflects cycles of indirect effects with different paths and varying lags. Differential productivity effects may exhibit a shorter lag time than the indirect consequences of negative externalities. And it is likely that these indirect effects, such as overurbanization, would take a significant time period to manifest their effects on economic growth.

TABLE 5

OLS ESTIMATES OF THE EFFECTS OF FOREIGN INVESTMENT
DEPENDENCE IN 1949 ON FOREIGN INVESTMENT DEPENDENCE
IN 1967

	ACCUMULATED FOREIGN STOCKS/GDP 1967		
	<i>b</i>	β	<i>t</i> -value
LDEB49GDP	37	55	3.55**
LGDP1950	-.18	-.22	-1.38
LGDPpc1950	33	.27	1.88
Constant	-5.76		-4.99**
Adjusted <i>R</i> ²47		

NOTE.—*N* = 33

** *P* < .01

It might be argued that these effects are spurious. There is a commonly held assumption that foreign capital is attracted to areas of relatively higher economic growth. It could be argued, therefore, that these findings actually represent the withdrawal of foreign capital and the subsequent decline in GDP growth resulting from this foreign capital flight.

However, the data do not support this argument. Bornschier and Chase-Dunn's (1985) PEN1 measure of capital penetration was regressed on my earlier measures of GDP per capita, calculated as change scores over the time period in question. GDP growth 1940–65 did not have a significant effect on the amount of foreign investment penetration in 1967.

CONCLUSION

The results of this study confirm that peripheral countries with relatively high dependence on foreign capital exhibit slower economic growth than those less dependent peripheral countries. These findings have been replicated using different measures of foreign investment dependence, GDP data, countries, time periods, and statistical methods. This is a significant and persistent negative effect, lasting for decades. Further, a structure of dependency is created that perpetuates these effects. The consequences of these effects, as described in the literature, are pervasive: unemployment, overurbanization, income inequality, and social unrest, to name a few. This is what all the fuss has been about.

The next step in this research should be a more complete specification of this model, to explicate the mechanisms by which these dependency effects occur. Foreign investment dependence also needs to be examined

from a world-systems perspective, a perspective that understands foreign investment dependence as a reflection of power relationships among countries that generates and maintains the core-periphery hierarchy in the world-economy. From this perspective, foreign investment dependence becomes a variable at the macro level of zones of the world-economy, rather than as a country-level variable.

APPENDIX

TABLE A1
CORRELATION MATRIX OF VARIOUS MEASURES OF FOREIGN
INVESTMENT DEPENDENCE

	FANGS38	DEBITS38	DEBITS49
<i>Listwise deletion.</i>			
DEBITS3888		
DEBITS4966	.75	
AFS6782	.82	.78
<i>Pairwise deletion:</i>			
DEBITS3888 (27)		
DEBITS4969 (44)	.75 (26)	
AFS6782 (44)	.87 (25)	.75 (52)

NOTE.—For listwise deletion, $N = 25$. Nos. in parentheses show no. of cases.
FANGS38 = foreign accumulated nongovernment stocks, 1938 (Lewis 1948),
DEBITS38 = debits on investment income, 1938 (World Bank 1949, 1951),
DEBITS49 = debits on investment income, 1949 (World Bank 1949, 1951),
AFS67 = accumulated foreign stocks, 1967 (Bornschier and Heintz 1979)

TABLE A2
COUNTRIES INCLUDED IN THE ANALYSES

Country	Table	Country	Table
Algeria	3	Lesotho	3
Argentina	1, 2, 3	Liberia	3
Australia	1	Madagascar	3
Austria	1	Malawi	3
Belgium	1	Malaysia	3
Bangladesh	3	Mauritania	3
Benin	3	Mauritius	3
Brazil	1, 2, 3	Mexico	1, 2, 3
Bulgaria	1	Morocco	3
Burkina Faso	3	Nepal	3
Burma (Myanmar)	2, 3	Netherlands	1
Cameroon	3	Nicaragua	2, 3
Canada	1	Nigeria	3
Chile	1, 2, 3	Norway	1
Colombia	1, 2, 3	Pakistan	2, 3
Costa Rica	3	Panama	2, 3
Cyprus	3	Paraguay	1, 2, 3
Czechoslovakia	1	Peru	2, 3
Denmark	1	Philippines	2, 3
Dominican Republic	2, 3	Portugal	2, 3
Ecuador	2, 3	Rwanda	3
Egypt	2, 3	Sierra Leone	3
El Salvador	2, 3	Singapore	3
Ethiopia	3	South Africa	2, 3
Finland	1	Spain	1, 3
France	1	Sri Lanka	2, 3
Gabon	3	Sudan	3
Ghana	3	Sweden	1
Greece	1, 2, 3	Switzerland	1
Guatemala	2, 3	Syria	3
Guyana	3	Taiwan	1, 2, 3
Haiti	2, 3	Thailand	1, 2, 3
Honduras	1, 2, 3	Togo	3
Hong Kong	3	Tunisia	3
India	1, 2, 3	Turkey	1, 2, 3
Indonesia	2, 3	United Kingdom	1
Ireland	2, 3	United States	1
Israel	2, 3	USSR	1
Italy	1, 2, 3	Uruguay	1, 2, 3
Ivory Coast	3	Venezuela	1, 2, 3
Jamaica	3	Yugoslavia	1
Japan	1, 2, 3	Zaire	3
Jordan	3	Zambia	3
Kenya	3	Zimbabwe	3
Korea, South	3		

REFERENCES

- Amin, Samir. 1974. *Accumulation on a World Scale: A Critique of the Theory of Under-Development*, 2 vols. New York: Monthly Review Press.
- Arbuckle, James. 1996. "Full Information Estimation in the Presence of Incomplete Data." In *Advanced Structural Equation Modeling*, edited by George Marcoulides and Randall Schumacker. Mahwah, N.J.: Lawrence Erlbaum Associates.
- Arrighi, Giovanni. 1994. *The Long Twentieth Century*. New York: Verso.
- Bornschier, Volker. 1980. "Multinational Corporation and Economic Growth: A Cross-National Test of the Decapitalization Thesis." *Journal of Development Economics* 7:191-210.
- Bornschier, Volker, and Thanh-Huyen Ballmer-Cao. 1979. "Income Inequality: A Cross-National Study of the Relationships between MNC Penetration, Dimensions of the Power Structure, and Income Distribution." *American Sociological Review* 44 (3): 487-506.
- Bornschier, Volker, and Christopher Chase-Dunn. 1985. *Transnational Corporations and Underdevelopment*. New York: Praeger.
- Bornschier, Volker, and Peter Heintz, eds. 1979. *Compendium of Data for World-Systems Analysis*. Zurich: University of Zurich, Sociological Institute.
- Boswell, Terry, and William J. Dixon. 1990. "Dependency and Rebellion: A Cross-National Analysis." *American Sociological Review* 55:540-59.
- Chase-Dunn, Christopher. 1975a. "The Effects of International Economic Dependence on Development and Inequality: A Cross-National Study." *American Sociological Review* 40:720-38.
- . 1975b. "International Economic Dependence in the World-System." Ph.D. dissertation. Stanford University, Department of Sociology.
- Dixon, William J., and Terry Boswell. 1996. "Dependency, Disarticulation, and Denominator Effects: Another Look at Foreign Capital Penetration." *American Journal of Sociology* 102 (2): 543-62.
- Firebaugh, Glenn. 1992. "Growth Effects of Foreign and Domestic Investment." *American Journal of Sociology* 98 (1): 105-30.
- . 1996. "Does Foreign Capital Harm Poor Nations? New Estimates Based on Dixon and Boswell's Measures of Capital Penetration." *American Journal of Sociology* 102 (2): 563-75.
- Frank, Andre Gunder. 1979. *Dependent Accumulation and Underdevelopment*. New York: Monthly Review Press.
- Grimes, Peter. 1996. "Economic Cycles and International Mobility in the World-System: 1790-1990." Ph.D. dissertation. Johns Hopkins University, Department of Sociology.
- IMF (International Monetary Fund). 1949. *Balance of Payments Yearbook*. Washington, D.C.: International Monetary Fund.
- . 1951. *Balance of Payments Yearbook*. Washington, D.C.: International Monetary Fund.
- Kentor, Jeffrey. 1981. "Structural Determinants of Peripheral Urbanization: The Effects of International Dependence." *American Sociological Review* 46:201-11.
- Kim, Jae-on, and James Curry. "The Treatment of Missing Data in Multivariate Analysis." *Sociological Methods and Research* 6:2.
- Korzeniewicz, Roberto Patricio, and Timothy Patrick Moran. 1997. "World-Economic Trends in the Distribution of Income, 1965-1992." *American Journal of Sociology* 102:1000-1039.
- Lewis, Cleona. 1948. *Debtor and Creditor Countries: 1938, 1944*. Washington, D.C.: Brookings Institution.
- Little, Roderick, and Donald Rubin. 1989. "The Analysis of Social Science Data with Missing Values." *Sociological Methods and Research* 18:2-3.

- London, Bruce. 1988. "Dependence, Distorted Development, and Fertility Trends in Non-Core Nations: A Structural Analysis of Cross-National Data." *American Sociological Review* 53:606-18.
- Maddison, Angus. 1995. *Monitoring the World Economy, 1820-1992*. Paris: OECD.
- Timberlake, Michael, and Jeffrey Kentor. 1983. "Economic Dependence, Overurbanization, and Economic Growth: A Study of Less Developed Countries." *Sociological Quarterly* 24:489-507.
- Wallerstein, Immanuel. 1974. *The Modern World System*. Vol. 1, *Capitalist Agriculture and the Origins of the European World-Economy in the Sixteenth Century*. New York: Academic Press.
- . 1979. *The Capitalist World Economy*. New York: Cambridge University Press.
- . 1980. *The Modern World System*. Vol. 2, *Mercantilism and the Consolidation of the European World-Economy, 1600-1750*. New York: Academic Press.
- . 1989. *The Modern World System*. Vol. 3, *The Second Era of Great Expansion of the Capitalist World-Economy, 1730-1840s*. New York: Academic Press.

Commentary and Debate

To conserve space for the publication of original contributions to scholarship, the comments in this section must be limited to brief critiques. They are expected to address specific errors or flaws in articles and reviews published in the *AJS*. Comments on articles are not to exceed 1,500 words, those on reviews 750 words. Longer or less narrowly focused critiques should be submitted as articles. Authors of articles and reviews are invited to reply to comments, keeping their replies to the length of the specific comment. The *AJS* does not publish commenters' rebuttals to authors' replies. We reserve the right to reject inappropriate or excessively minor comments.

RESPONSE TO HIRSCH'S REVIEW ESSAY

Paul Hirsch has written a long and detailed essay on what I took to be a succinct overview of the state and direction of institutional theory ("Sociology without Social Structure: Neoinstitutional Theory Meets Brave New World" [*AJS* 102 (1997): 1693–1701], review essay on *Institutions and Organizations: Theory and Research* [Sage, 1995]). While he sprinkles faint praise throughout his review, Hirsch seriously distorts some of the most important ideas and arguments of my book. Of many possible issues, I select three as primary.

Hirsch treats my analytical schema as though it were intended as "a forced-choice conceptual dichotomy" (p. 1704)—there being three categories! He reports that I insist that all institutions must be forced into one and only one category. I employ the word "analytic" in what I believe to be a conventional way. The purpose of the schema is to enable investigators to disentangle the strands of complex concrete phenomena into categories for purposes of analysis. It is true that I attempt to illustrate my typology by categorizing selected studies into one or another of the types defined, but I stress throughout my review that I do so on the basis of the primary emphasis of the researcher. I early and often insist that all but the most limited institutions are made up of combinations of regulative, normative, and cognitive elements.

Second, Hirsch indicates that I show a distinct preference for the "cogni-

Permission to reprint a comment printed in this section may be obtained only from the author.

tive" strand of institutional analysis. Indeed, I indicate on page xv of the introduction that "my major interest is in explicating the sociological variant" of the "new institutionalism." After all, these newer contributions are less well developed and the least well understood. In paying more attention to the newer, cognitive work on institutions, Hirsch suggests that I slight and belittle the contributions of the other, earlier formulations. I do not believe that this is the case. I do make the apparently unforgivable error of referring to normative approaches by sociologists to the study of institutions as "traditional"—since they have been carried out for over 100 years—and this sounds to Hirsch like a serious dismissal of the work that he and many other contemporary sociologists do. It is not intended as such.

Hirsch points out that each of the conceptual frames directs attention to some important issues, but also "directs attention away" from other problems. Precisely! That is what a good set of analytical concepts does. To take a more cognitive perspective is to raise questions about where organizational templates or models come from that are not addressed by regulative or normative conceptions. To take a more normative perspective is to raise questions about social roles and obligations that are slighted by the cognitive and regulative models, and so on.

Finally, because I point out that the function of institutions is to secure stability and meaning, Hirsch concludes that in my "brave new world, compliance is the goal shared by all institutional theories" (p. 1702). Here, my descriptive analysis is misread as prescriptive or normative injunction. Hirsch's criticism is particularly unjust because, more so than most neoinstitutionalists, I emphasize the ways in which organizations respond to institutional pressures in ways other than compliance—by avoidance, defiance, manipulation, and other resistant tactics (see pp. 124–30 in *Institutions and Organizations*).

By crafting his lengthy review essay, Paul Hirsch has paid me the compliment of taking my book seriously. I just wish he had read the book more objectively.

W. RICHARD SCOTT

Stanford University

REJOINDER TO SCOTT

A way of seeing is also a way of not seeing (Burke)

Even a hermit in bleakest Antarctica must be aware of the organizational science paradigm war by now. (McKelvey)

I am pleased to see how much Scott concedes I got right. For those points in the essay he elects to respond to, Scott *agrees* that in *Institutions and Organizations*, albeit analytically, he reduces the "function of institutions" to a command and control mission: "to *secure* stability and meaning" (emphasis added). Scott also *agrees* that he directs attention away from organizational sociology's focus on organizations as actors and the discipline's "traditional" [*sic*] study of contemporary social issues as contests over the social definition of roles, norms, and obligations. More important, *we further agree* that his book shifts attention away from these latter concerns and problematics to what Scott calls the "new" and more "cognitive" perspective on organizations.

In the commentary and book, Scott bases this open effort to alter the research direction of the field on the peculiar assertion that the ideas and philosophies characterizing the third ("cognitive") pillar in the book's typology warrant disproportionate attention because they are "new" and original. What he does not address in his comment is my essay's challenge to this position and the two directly contrary observations I provided to support it. The first noted that a comparison of the publication dates of the "recent" works cited to indicate the emergence of "new" institutionalism finds little difference between them and the publication dates of works embodying the alternative (regulative or normative) points of view given the less attractive label "traditional." None of the positions in or viewpoints that help to define each pillar has been around for any longer or shorter time than the others. Also not addressed is a second observation in the essay: if one substitutes the terms "grand theory" for what Scott calls "new institutional" and "middle range" for what he has labeled "traditional," it is further apparent that the topics under discussion here as "new institutional" are neither new nor current for any of the social sciences. In sociology, they date back to the classic critiques and debates over research agendas for the discipline, between Parsons and Merton, and others before them. (I also noted that, in the first part of the book, and oddly decoupled from what follows, Scott's text also encourages this observation.)

If the basis for asserting distinctions between "old," "new," "traditional," "neo," and so on, is lacking in substance, what else might help explain (per McKelvey 1997) the representations so ably articulated by Scott to so pejoratively label the alternative perspectives? Much of the debate over

this book is around which of the many competing perspectives on, and interpretations of, terms like "institutional," "social construction," and "process" will prove necessary for the success of articles and book manuscripts submitted for publication and dissertation topics in the larger discipline and its subfields.

I open my essay with a credit to Scott: "he is more open to such alternative perspectives than a literal reading of *Institutions and Organizations* may suggest" (p. 1703). The empirical question at hand is to decide how comfortable *the book* leaves the reader about the prospect of doing research in the *alternative* (normative or regulative) "nongenerative" perspectives. Unfortunately, in my teaching experience and conversations with doctoral students, when asked if Scott's presentation is sufficiently neutral to make research in each pillar appear equally interesting and protracted, students answer with a resounding no.

One purpose of my essay was to flag that Scott's recent work in this arena reveals an interesting and new, more activist persona, that is unlike the more neutral stance in his classic texts on the field of organizational sociology (p. 1704). If we agree on as much about the book's messages as I pointed out above, Scott's last "defenses"—that I did not read this political book "objectively," or that (when I did) he did not intend for certain conclusions to be drawn—are hardly relevant to its wider interpretation and reception.

Certainly I did not assert that in this work Scott rules out the possibilities and potentials for studying institutional change and organizational action. As he notes, they do indeed eventually make appearances in this book. On the pages he notes (124–30), these words may be found; however, Scott's resistance still seems clear when nearly every statement about their importance is preceded, or qualified by, additional comments that consistently dilute the value accorded to taking change seriously.¹ If the book also suggests that empirical work on institutions combines the elements of all three pillars, my reading still finds that this, too, is stated far less clearly and frequently than Scott suggests in his comment.

Since, as Scott also notes, these can remain matters of interpretation, the empirical question of whether readers come away from the book thinking the study of dynamics, contests, and changes are the way to go, remains open. My reading, however, suggests that the message remains

¹ As an aside, I should clarify that the "forced-choice conceptual dichotomy" I noted in the essay was Scott's clear distinction/choice offered between the "realist" ontology, subsuming the regulative and normative perspectives, and the "social constructionist" ontology, into which he codes the cognitive perspective. Scott presents these as (the two) opposing ontologies. This is discussed in a different portion of my essay than Scott's typology of three institutional pillars. It is the latter three categories which I argued Scott does not clearly enough permit research to cross or combine.

clearly tilted toward stasis, stability, and grand assumptions and away from examining important questions about conflict, mechanisms, and contemporary social issues.

PAUL HIRSCH

Northwestern University

REFERENCES

- Burke, Kenneth. 1935. *Permanence and Change*. New York: New Republic.
McKelvey, Bill. 1997. "Quasi-Natural Organization Science." *Organisation Science* 8 (4): 352–380.

Review Essay

AN INTELLECTUAL'S STOCK IN THE FACTORY OF MEXICO'S RUINS¹

Mexico: Biography of Power by Enrique Krauze. Translated by Hank Heifetz. New York: Harper Collins, 1997. Pp. 872. \$35.00.

Claudio Lomnitz
University of Chicago

At the end of every presidential term (*sexenio*), Mexican presidents become involved in a frenetic race of inaugurations; their posterity depends on it. Hospitals, museums, universities, dams, highways, subways—all of the signs of the modernization and progress that every president promises—must be inaugurated, along with a large bronze plaque giving credit to the president, whether the building is finished or not. My brother, a scientist, once witnessed the inauguration of a research facility by the outgoing president López Portillo in 1981. The inauguration occurred in a building that was made to look finished, complete with lawns, potted plants, and the rest of it. As soon as the president left, a presidential team came in, rolled up the grass, picked up the potted plants, and took them to the site of the next inauguration.

This practice, which betrays so much about the economy and legitimacy of Mexican presidentialism, is certainly one of the sources of what Brazilian literary critic Beatriz Jaguaribe has called “modernist ruins.” The rush to legitimize a presidency or a governorship is enmeshed with the economy of public expenditure, and both conspire to produce veritable monuments to the grandiloquence and corruption of the governing elites that are, at the same time, inhospitable and alienating for the intended user (“the public”). The fascinating thing about these modernist ruins is that they betray the gestural quality of much of Mexico’s state-led modernity. The central tenet of architectural modernism (utility, practicality) serves as a screen for a second rationale, which is political: the story of Mexico’s progressive state veils an enormous pork barrel.

This aspect of Mexico’s modernity was most poetically captured by the Scottish eccentric and surrealist Edward James, who built majestic

¹ Correspondence may be directed to Claudio Lomnitz, Department of History, University of Chicago, 1126 East Fifty-ninth Street, Chicago, Illinois 60637

cement ruins for the jungles of the Huasteca to swallow up. When he was asked why he built this costly extravagance, Edwards claimed that it was to confuse the archaeologists of the future. Like James's ruins, Mexico's "modernist ruins" have very personal signatures, which are often those of the president who sponsored them. So, whereas archaeologists of the pre-Columbian past use site names to label historical epochs (e.g., Monte Albán I, II, III, or Tlatilco IV, V, and VI), archaeologists of Mexico's modernist ruins would be wise to rely on the names of the presidents who sponsored them, for example, Alemán I and II or López Portillo I, II, and III.

Although the discussion of modernist ruins usually brings to mind housing projects, hospitals, bridges, and basketball courts, Mexico's cultural world is also littered with these ruins. The central axis of cultural modernity—which is a productive relationship between science, art, and the constant improvement of the quality of life ("progress")—was historically so feeble in Mexico that, beginning in the 1920s, the state adopted an aggressive, active role in strengthening it. This role has been as open to demagoguery and corruption as any other modernizing project.

Until the 1960s, the state's function as patron of the sciences and the arts had met with relative success—the National Autonomous University was built, as was the National Polytechnical Institute. The arts flourished under state patronage, and Mexico began to make a credible bid for a place among modern nations. The revolutionary prestige of the government and the accelerated modernization that began around 1940 fostered a relatively snug relationship between middle-class ideals of mobility and the state's self-image as the prime engine of modernization.

Mexican sociologist Ricardo Pozas (1993) has shown how this relationship first cracked in 1964, when medical students and young doctors rejected the state's authoritarian forms of decision making and embarked on a series of strikes that were violently suppressed. The sequel and culmination of this conflict occurred in the student movement of 1968, which ended with the massacre of hundreds of students at Tlatelolco Square in Mexico City.

The killings at Tlatelolco provoked a new spurt of construction of modernist ruins. Under President Echeverría the whole of Mexico's university system expanded far beyond the country's capacities, which meant filling the university with a staff that was not always well qualified. Although the results of this huge expansion of the educational system in the seventies were mixed, criticisms of its perverse effects were particularly harsh because the formula of state-driven expansion was no longer sustainable after the state's fiscal crisis in 1982.

The National University and other public institutions came under severe scrutiny, and their ruinous aspect was widely publicized as the De la Madrid administration slashed its support of Mexican public institutions of higher learning. This was the dawn of a new era in Mexican cultural life, an era marked by privatization and by growing differences between an increasingly proletarianized mass of low-prestige teachers, a

somewhat fancier stratum of publishing academics, and a new cultural elite that fuses writing with business.

Changes in Mexico's cultural world have been deep, yet the analysis of their impact on the quality of cultural production has been suspended to a surprising degree. There is much that is new in the institutional arrangement of Mexican cultural life since the eighties, including changes in training programs and in the profile that is expected for entering a university career, growth of private and public universities, and the emergence of cultural groups with wide media access.

There are signs, however, that the time is ripe for a critical look at today's cultural milieu, for the era's first monumental modernist ruins are now becoming clearly visible. This year it seems that Mexico City's main private art museum may close its doors or at least drastically reduce acquisition. It is also clear that most Mexican private universities are not funding much research. However, in the world of culture, the most significant ruins are always the cultural works themselves. The appearance of Enrique Krauze's *Mexico: Biography of Power* is a landmark in this respect; it is a "period piece" that allows us to scrutinize the effects of power on intellectual production in a sector of Mexico's intelligentsia.

I propose to do just this. A discussion of the organization of Krauze's book, of the connections between Krauze's intellectual project and his position in Mexico's cultural milieu, and an appraisal of the value of this book as a work of history open a clear perspective on the use of history as a gesture in the struggle over who gets to represent Mexico.

Organization

Mexico: Biography of Power is Enrique Krauze's most ambitious book. It combines into a single work three books originally published in Spanish (*Biografías del poder*, about the leadership of the Mexican Revolution; *Siglo de caudillos*, about the Mexican presidency in the 19th century; and *La presidencia imperial*, which covers the Mexican presidency from 1940 to the present). In addition, this English-language volume offers a brief synthesis of political power and political culture in the colonial period. This is the only single work available, in English or in Spanish, that covers such vast territory.

The complexity of the subject matter is made manageable by giving history a direction and a premise. Both of these are offered with disarming simplicity. For Krauze, the history of Mexico is the history of the struggle for democracy. So much so that, echoing Fukuyama, he ends this book by asking Mexicans "to bury once and forever Cuauhtémoc with Cortés, Hidalgo and Iturbide, Morelos and Santa Anna, Juárez with Maximilian, Porfirio with Madero, Zapata and Carranza, Villa and Obregón, Calles with Cárdenas, all of them reconciled within the same tomb. But Mexico would have to be less pious toward its modern actors. There can be no reconciliation with Tlatelolco" (p. 797).

Krauze feels that the 1968 massacre at Tlatelolco should not be forgotten because that conflict was largely about governmental democracy. However, the fact that the 1968 movement did not involve or affect Mexico's peasants nor the majority of its poor does not seem to matter: Mexico's peasants are asked to "bury Zapata," who called for land for those who work it, but never to forget a middle-class movement that demanded democracy. The organization of political history around the story of democracy is highly problematic in a country whose fundamental viability was in question during most of the 19th century. Moreover, although democracy has been a significant political issue during most of Mexico's modern history, it has often not been the principle political aim or site of contention.

For instance, the Mexican Revolution (1910–20) begins as a democratic revolt under Madero, but it quickly turns into a broad-based and rather inchoate social revolution with variegated demands ranging from agrarian reform, to labor laws, to national control over resources, to radical state secularism. On the whole it is fair to say that these demands, and the dynamics of the struggle for power itself, overshadowed democracy as the main issue. This fact is confirmed in the political success of the official state party (PRI), a party that was deeply undemocratic but that left considerable room for social demands. In short, although the organization of Mexico's political history around the epic of democracy is pleasing for American readers and to some political groups in Mexico it is not defensible as the key to understanding that history.

The book's central premise shares the pleasing simplicity of its teleology: "This book threads the lives of the most important leaders during the last two centuries into a single biography of power, but I am in no way subscribing to an outmoded (and unacceptable) great-man theory of history" (p. xv). Thus, while writers and academics the world over worry about the "death of the subject," Krauze is busy anthropomorphizing national history and providing it with a "biography":

What I hope to convey is that in Mexico the lives of these men do more than represent the complexities and contradiction of the country they came to govern or in which they took center stage for a time at the head of armies fighting for change or for a return to the past (or for both). The accidents of their individual lives also had an enormous effect on the directions taken by the nation as a whole. Personal characteristics and events that in a moderately democratic country might be mere anecdotes—interesting, amusing, or trivial—can in Mexico acquire unsuspected dimensions and significance. An early psychological frustration, a physical defect, a family drama, a confused prejudice, a tilt one way or the other in a man's religious feelings or his passions, even a local tradition automatically accepted could literally alter the fate of Mexico, for better or for worse. (P. xv)

According to Krauze, then, presidential biographies in Mexico collectively shape what he mystically calls the nation's "biography of power." However, he does not want this to be identified with a "great-man theory of history" but wishes instead to provide the premise with a kind of cul-

tural specificity. This is because Mexico's historical roots combine "two traditions of absolute power—one emanating from the gods and the other from God—[he means the Aztec and the Spanish traditions] this political *mestisaje* conferred a unique connection with the sacred on Mexico's succession of rulers" (p. xiv). What we have, then, is a great-man theory of history whose validity is confined to Mexico.

As a result, Krauze continually asserts that Mexico is unique and fundamentally different from the rest of the world. This exceptionalism is convenient because it allows him to ignore the parallels between Mexican history and other histories, parallels that would diminish the force of the contention that presidential biographies have systematically "altered the fate of Mexico." On the other hand, since Krauze claims exception for Mexico on the basis of the peculiarities of the Aztec and Spanish mixture, this leads straight back to Mexico's official history, which this book distinctly reproduces: Martín Cortés (son of Hernán Cortés and La Malinche) was "the first Mexican" (p. 52); Hernán Cortés was "the spiritual antithesis" of Moctezuma (p. 44); Moctezuma and Cortés "created a new nationality the instant they met" (p. 47); there was no "true ethnic hatred" in Mexico from the colonial period forward (p. 49); slavery in Mexico was sweeter than in the United States (p. 50), and so on. In short, the fabricated saga of the mestizo as national protagonist is swallowed whole—hook, line, and sinker. The authoritative narration of Mexico's fate and fortune rehearses and reaffirms official history, but with a twist: instead of culminating with the progress wrought by the Mexican Revolution (which had been the End of History until recently), it culminates with the democracy that Krauze's 1968 generation is supposed to have engendered.

Krauze: Biography of Power

Krauze's history can be read in two keys: the first key is the saga of democracy into which he wants to shoehorn Mexican political history, the second is the saga of his own intellectual genealogy. This second epic, which is barely visible to an English-speaking audience, is nonetheless critical, because Krauze is in the business of representing the nation to the outside, trying hard to garner credentials with which to construct himself as the kind of privileged interlocutor that other Mexican intellectuals have been: Octavio Paz, Carlos Fuentes, Diego Rivera, Rufino Tamayo.

Enrique Krauze began his career with a book on what he called "intellectual *caudillos*" of the Mexican Revolution. (The term *caudillo* originally referred to military leaders whose charisma allowed them to vie for control over countries and regions. It is a political form that was characteristic of Spanish America's 19th century). Krauze then hitched his wagon to the star of Daniel Cosío Villegas, a prominent liberal historian who directed El Colegio de México and who created a workshop that

was known as the "factory of Mexican history," where much of the history of the Porfiriato and the Mexican Revolution was written.² After Cosío's death, Krauze became the impresario and subdirector of Octavio Paz's cultural magazine, *Vuelta*, and it was from this publication that he derived most of his intellectual cachet.

In an effort to create a voice for himself, and perhaps to emerge from under the long shadow of his mentors, Krauze identifies as a member of the 1968 generation, a generation that was marked by the student movement and by its violent end at the hands of the Mexican state. Like a number of others, Krauze relies on this identity to acquire the semblance of purity. He sees himself as a liberal and even as a "heretic" (see Krauze 1992), an independent intellectual who criticizes Mexican authoritarianism from the sanctity of his private world.

In fact, however, Krauze's prestige and cultural power do not come from 1968, nor is he comparable on an intellectual plane to Cosío Villegas, let alone to Octavio Paz. Krauze's prominence is, instead, an effect of a more recent story. With the debt crisis in 1982, the Mexican government came down hard on all salary earners, real minimum wages plummeted to half in under five years (a fact that, like almost every economic consideration, goes unnoted in Krauze's book). Among the wage-earning population, one of the sectors that was hit hardest was the educational sector, and the universities in particular. The government was unwilling to maintain university salaries at their traditional middle-class levels, and so it created a system of evaluation that sidestepped university promotion regulations and rewarded only productive academics. "Publish or perish" came to have a very literal meaning in the Mexican academy. However, the process of internal stratification in the university system did not come without a substantial cost both for the prestige of academic work and for the possibility of surviving as a young scholar. As a result, whole generations of potential scholars were either significantly slowed down or destroyed.

At the same time that the Mexican state strangled its universities, it did not abandon its patronage and contact with intellectuals. The De la Madrid (1982–88) and Salinas (1988–94) governments coupled their tight policies toward the university with generous contracts and subsidies to specific intellectual groups. The principle groups gravitated around two literary/political journals: *Vuelta* and *Nexos*. These two groups accumulated vast cultural power in the 1980s and 1990s: Hector Aguilar Camín, former director of *Nexos* and member of the 1968 generation and erst-

² In fact, the central thesis of *Mexico: Biography of Power* (i.e., the preponderance of the president's biography over Mexican history) is derived from an essay by Cosío Villegas that was written against President Echeverría—a president that had an especially strong delusion of omnipotence—titled *El Estilo personal de gobernar* (1974). The theme of that essay, which was that in Mexico the president's personal whims had become a kind of *raison d'état*, is magnified by Krauze into *the* key to the whole of Mexican history.

while leftist, was a close friend to Carlos Salinas de Gortari. He created a publishing house, Cal y Arena, whose books were widely distributed publicized by *Nexos*-controlled public television Channel 22.

On his side, Enrique Krauze, the principle entrepreneur of the *Vuelta* group, received support from president De la Madrid for his "biographies of power" project (comprising the Porfiriato to Cárdenas sections of *Mexico: Biography of Power*), which was subsequently printed by the government-owned publishing house Fondo de Cultura Económica, a prestigious press that side-stepped its traditional role of publishing scholarly work.

During that same period, Krauze and *Vuelta* began doing business with Televisa, Mexico's television giant that had, thanks to its special ties to the government, effectively been a communications monopoly for decades. Televisa had a largely negative role in Mexico's transition to democracy, a fact that has been widely recognized by independent political observers of Mexico, including the United Nations. This did not stop self-styled democratic hero Enrique Krauze from becoming one of the company's partners. Krauze is coowner of Clio, a publishing house devoted to popularizing his version of Mexican history and producer of historical soap operas that have devoted some effort to rehabilitating Porfirio Díaz (1876–1910), the liberal dictator and former archvillain of official history.

In short, Krauze's power was amassed during a period when the government turned its back on public education and research and subsidized a process of cultural privatization that had similar characteristics to other privatizations: enormous concentration of power in very few hands and the formation of a new elite.

Whereas Daniel Cosío Villegas's "factory of history" was built in a public institution, Krauze's factory of history is private. For big rollers in Mexico's cultural enterprises, research is a menial task. Thus, where most historians work alone or with one or two assistants, Krauze lists 16 in his acknowledgments, two of whom are as accomplished as historians as Krauze himself.³ His heavy reliance on this private factory is the reason why this book is such a good mirror of presidential power: the resources that Krauze musters have allowed him to write a monumentally ambitious work, but his methods make him unsure at every turn. *Mexico: Biography of Power* is a hollow monument.

Krauze as Historian

This book's main empirical contribution is a set of interviews that the author or his assistants made with important political figures as well as a much-publicized, but rather disappointing, diary of president Díaz Ordaz. Most of the book, however, is based on published documents as well

³ These are Margrita de Orellana and Aurelio de los Reyes.

as on secondary sources. The use of these secondary sources provides another key for the archaeologist of Mexico's modernist ruins.

During the past 20 years or so, U.S. and British historians have written a sizable proportion of the most relevant works on Mexican history, yet the work of historians such as John Coatsworth, Alan Knight, Eric Van Young, Gilbert Joseph, Anthony Pagden, John Tutino, Florencia Mallon, and Stephen Haber is not cited, nor—in most cases—are their ideas assimilated in the text, despite their indisputable relevance to the subjects covered. Like the politicians who have always stressed Mexican exceptionalism, Krauze too is interested in Mexico's insularity. By turning his own coterie of friends and mentors into the principle thinkers and actors in Mexican history he can easily aspire to become Mexico's representative in the media.

The use of the work of Mexican scholars is equally problematic. For instance, in his treatment of the 1968 movement, a chapter that is meant to be the high point of the book, Krauze gives preeminence to two intellectuals—Cosío Villegas and Octavio Paz—both of whom were marginal to the movement and of an older generation but were nonetheless central to Krauze's own development. Cosío Villegas gets no less than 33 mentions in the text of this book; the Mexican historian Edmundo O'Gorman, who was arguably a more profound thinker, gets none. Perhaps the oversight is due to the fact that O'Gorman publicly disapproved of Krauze's biographies of power. Citations of significant books written by members of a younger generation of Mexican scholars are another notable absence, perhaps because they are potential competition.

In addition to the political motives of these oversights, there is another likely cause for Krauze's sloppy use of secondary sources: the factory. This hypothesis comes to mind because there are a number of instances when a key historical work is indeed cited, but its conclusions are not assimilated in the analysis. Or else a work is cited in one context (perhaps being worked on by one of his research assistants) but then fails to appear as a source in another part of the book where its inclusion would have been particularly germane.

For example, French historian François Xavier Guerra has developed quite a complex view of the modernization of the Mexican state in the 19th century. Guerra's view is that between Independence (1821) and the Revolution (1910), Mexican political society changed from being made up of corporations that were built around personal ties in villages, guilds, and haciendas, to a modern society in which these personal ties could no longer hold the country together. As a result, the personal power of Porfirio Díaz (1876–1910) is, for Guerra, both the culmination and the swan song of what Krauze calls a "biography of power." Guerra is cited on a factual matter, but his general argument is ignored. Moreover, Guerra fails to appear in Krauze's discussion of political theory of independence, where he would have been very helpful. In sum, the cavalier use of secondary sources is possibly the only true sense in which Krauze can be called liberal.

The Authority of Opinion

Enrique Krauze has had two principle mentors, Daniel Cosío Villegas and Octavio Paz. As shown above, Krauze took Cosío Villegas's factory of history, privatized it, and made it into his own political machine. From Paz, Krauze has tried to emulate grandeur, scope, and boldness. The result is not always bad. *Mexico: Biography of Power* is certainly a readable book. However, Krauze's attempts at Paz-like boldness also have a very perverse effect in that they liberate this book from the usual strictures of historical evidence.

Krauze has made a name for himself in Mexico by calling for a "democracy without adjectives," but he seems entirely incapable of offering a history without opinions.⁴ More often than not, these opinions are stated as if they were facts. In *Mexico: Biography of Power* we are asked to believe, for instance, that there were only two "true ethnic wars" in Mexican history (p. 780) and that Cosío Villegas's criticisms of President Echeverría (1970–76) were the bravest thing any Mexican had published in 100 years (p. 746); we also learn that "Juárez the Indian" "was all religion" (p. 167) and that his invocations of God and Providence were carried out "without hypocrisy" (p. 166). In short, the reader is exposed to the dictatorship of what might usefully be labeled "the Krauzometer."

The translator, Hank Heifetz, has done a commendable job not only in avoiding the annoying changes of register that characterize Krauze's Spanish prose, but also in trying to tone down the Krauzometer as much as possible. So, for instance, in *La Presidencia imperial* (the Spanish-language book that comprises parts 4 and 5 of *Mexico: Biography of Power* and that appeared simultaneously with it; Krauze 1997), Octavio Paz's *Labyrinth of Solitude* is "the most important book of the Mexican twentieth century" (certainly a "100" on the Krauzometer; see p. 152), but it is only "one of the most important books of the Mexican twentieth century" in the English-language translation (only an "80" on the Krauzometer; see p. 364). Similarly, in Spanish, Krauze asserts boldly that president Díaz Ordaz (1964–70) did not lie in his memoirs (p. 355); in English he asserts that "it is unlikely that they are all lies" (pp. 728–29). In Spanish, Miguel de la Madrid won his election because the people voted for him personally and not for the PRI (p. 402); in English the people voted not for De la Madrid personally, but rather for his platform of moral renovation (p. 763). Moreover, in Spanish, De la Madrid won the election with 76% of the vote (p. 402), whereas in English he seems only to have received 68% (p. 763). In this book opinions are facts, and both change along with the intended readership.

⁴ See Krauze's *Por una democracia sin adjetivos* (1986). Not surprisingly, the phrase "democracy without adjectives" does not belong to Krauze, but is instead Rafael Segovia's (1980).

Biography and Power

Certainly, Krauze's factory has produced a readable book, with some new information and a wealth of anecdotes. Although none of this information makes a significant mark in the historical interpretation of modern Mexico, it does add richness and legibility to this facile and ideologically loaded text. In Mexico, Krauze's version of history is being massively consumed in soap operas, which is an appropriate—though perhaps not harmless—venue for it.

There is, in addition, another good selling point for this book, which is the idea that biography is a useful vantage point for political analysis. I have already argued that this interest in biography led Krauze to the great-man view of history that he allegedly rejects, but more attention to Krauze's biographies is warranted.

The first thing to note about these presidential biographies is that they rarely provide the kind of psychological insight that the author was hoping for. This unevenness is due not only to the space and detail devoted to various presidents (Miguel Aleman gets 75 pages, Manuel Avila Camacho gets 27, Miguel de la Madrid gets 8) but also to the format of the chapters. For instance, whereas we get an attempt to portray the family history and youth of presidents and caudillos between Porfirio Díaz and Gustavo Díaz Ordaz (1876–1970), there is no parallel information for the more contemporary presidents (beginning with Echeverría). Krauze thereby declines any attempt to provide a more profound portrait of the three presidents with whom he has had a personal relationship (De la Madrid, Salinas, and Zedillo).

The irregularity of the quality of biographical insights is also a product of Krauze's rush to represent, which leads inevitably to an imprudent reliance on common sense. For instance, Krauze tells us that "revolutions have been organized around ideas or ideals: liberty, equality, nationalism, socialism. The Mexican Revolution is an exception because, primordially, it was organized around personages. . . . The local histories from which they [these personages] began, their family conflicts, their lives before rising to power, their most intimate passions—all are factors that might have been merely personal, though perhaps representative, if these were merely private lives. But they could not be in Mexico, a country where the concentration of power into a single person (tlatoani, monarch, viceroy, emperor, President, caudillo, *jefe*, *estadista*) had been the historic norm across the centuries" (pp. 243–44). The trouble with this statement is that no distinctions are made regarding the significance of biographies, say, for a tlatoani and for a president, or for a caudillo and a monarch. Instead of attempting to specify these different forms of power, and then seeing their connection to biography, they are constantly collapsed into a single composite, which is then—sometimes anachronistically—turned into the quintessence of Mexican-ness.

Throughout the book terms such as "monarch," "tlatoani," "theocratic," and "caudillo" are used as metaphors for other forms of power. The Mexi-

can presidency is "like a monarchy." The president is "like a tlatoani." Presidential power is "almost theocratic." José Vasconcelos and Daniel Cosío Villegas were intellectual "caudillos," and communications magnate Emilio Azcárraga was a "caudillo" of industry. These comparisons and metaphors may be innocent enough in daily parlance, but if your thesis is that there is a special connection between the details of a leader's biography and the country's destiny (p. xv), then the difference between an actual monarchy and something that has similarities to a monarchy, an actual caudillo and someone who is compared to a caudillo, an actual tlatoani and a president, becomes critical.

For example, the power of a revolutionary caudillo like Emiliano Zapata was, especially in its origins, charismatic. People followed him because they shared his cause, were often in desperate straits, and because they believed in him. Zapata's biography is critically important because it is the source of the social connections of his inner circle (whose biographies in turn affect outer circles), and because his persona gave credibility and direction to the revolutionary movement as a whole. As a result, the epic of Zapata's life takes a messianic turn, similar to what we find in a number of revolutionary caudillos in Mexico, beginning with Miguel Hidalgo, whose political usage of the passion play was perceptively analyzed by Victor Turner (also not cited by the author).

Krauze argues that the biography of a Zapata and of a Hidalgo is critical for understanding their movement's destinies, but one might argue, conversely, that the construction of their personas was shaped by the context in which they arose as leaders. It is certainly no biographical accident that led Zapata, Hidalgo, and even Madero to take up a messianic Christian narrative and construct their persons around it. Specific forms of power such as presidencies, monarchies, grassroots leadership, and so on, imply different kinds of relationships between the leader's biography and the exercise of power. For instance, in European monarchies, the idea of "the king's two bodies" implied full identity between the king's well-being and the prosperity of the land. The king was an embodiment of his kingdom. Indeed, in the case of Spanish America, Philip II decreed the production of censuses and maps of the entire realm (the famous *Relaciones Geográficas*). The maps and descriptions he received were concentrated in his palace at El Escorial and in the office of the Royal Cosmographer, and the information in those censuses and maps was privy to the king. At the same time as he received the maps, he sent out portraits of his person to the four corners of the realm: the king concentrated the full image of the realm in his palace, the realm received, in its stead, the bodily image of the king (see Mundy 1996).

The relationship between biography and the application of power in this case is certainly distinct from that of Mexico's 19th-century presidents. The connection between presidential power and personal benefit inverted the central dogma of monarchy. Nineteenth-century caudillos like José María Morelos and even like Santa Anna wanted to be thought of as *servants*, not lords, of the nation. As a result, 19th-century presidents

("caudillos") routinely modeled their public personae after Cincinnatus—a renouncer (much like George Washington in the United States and Rosas in Argentina). However, Krauze wrongly reduces Santa Anna's constant show of retreating from the presidential chair to a psychological quirk ("He detested the direct and daily exercise of power"), when in fact it was a variation of a classical theme in the theater of presidential power in 19th-century Spanish America.⁵ Whereas the monarch identified his personal welfare and prosperity with that of the realm, early presidents and revolutionary caudillos used personal sacrifice as a legitimating device. As the presidency became a stable political institution, the office began to require less dramatic personal sacrifices and the image of the "civil servant" became more prominent—this was the image that Juárez adopted for himself, but it was not routinized in Mexico until well into the 20th century.

Krauze ignores all of this. For him, charismatic power is a constant in Mexican history, the product of a mythified fusion of Aztec and Spanish "theocracies." As a result, he reduces the differences in the persona of various leaders to the details of their biographies. This error leads to the kind of Mexican exceptionalism that I objected to earlier (to the proposition that there is something about Mexico that makes all of its leaders into tlatoanis—or did, until the fateful events of 1968, which brought about a new generation, led by Krauze among others, who have finally brought democracy to Mexico, the End of History). It also leads him to curious attempts to differentiate "authentic" from "inauthentic" leaders.

Antonio Lopez de Santa Anna for Krauze is the epitome of the fake. His power was theatrical, operatic, and worse, it was divorced from the nation's roots—never mind that "the nation" did not yet effectively exist. Thus, commenting on the rise of Benito Juárez (who, unlike Santa Anna, is portrayed here as being 100% authentic—"a pure-blooded Indian"), Krauze says that "the country would now be governed by a group of young mestizos who were closer to Mexican soil, closer to indigenous roots" (p. 151). This statement brings us back to the fundamental characteristic of this ruin: it is little more than a reenactment of the national myth for the nineties.

Conclusion

In *The Critique of the Pyramid*, a post-1968 reflection on what had gone awry in Mexico, Octavio Paz wrote a trenchant criticism of Mexico's National Museum of Anthropology. His main complaint was that the architecture of the building and its layout made the museum's Aztec Salon

⁵ See, e.g., my own book, *Exits from the Labyrinth: Culture and Ideology in Mexican National Space* (Lomnitz 1992, pt. 2, chap. 2). For Argentina, see Jorge Myers (1995); other Latin American illustrations can be found in Lynch (1992).

into the culmination and synthesis of all pre-Hispanic culture. This construction of the Aztec empire as both the centerpiece of the pre-Hispanic world and the antecedent of the independent Mexican nation negated cultural pluralism, idealized a strong central state, and falsified the pre-Columbian past.

Krauze's book is very much like that museum. The fusion and confusion of *tlatoanis*, *caudillos*, viceroys, and presidents, and the thesis that the course of Mexican history was dictated by Díaz Ordaz's ugliness, by Santa Anna's theatricality, and by Juárez's religiosity and purity makes this book as much of a Mexico City-centered account of the history of power in Mexico as the Museum of Anthropology ever was. In this allegedly critical review of the Mexican presidency, the presidents are fetishized, and the social history of the country is collapsed into nationalist myth.

The peculiarity of Krauze's generation of mythmakers is that they are not builders of state institutions, but have instead used state patronage to build private niches for themselves. Two Mexican intellectuals of the 1968 generation have been emblematic in this transition: Hector Aguilar Camín (former editor of *Nexos*) and Enrique Krauze (former subdirector of *Vuelta*). These intellectuals have been in the business of creating their own "factories of culture." They now speak from these niches and ventriloquize "civil society," much as Maya priests once interpreted the commands of a talking cross.

So far this new mode of cultural production has counted on the support of the Mexican state, of some powerful (government-related) businessmen and, by now, on its own private resources. The system also benefits, however, from the fact that the readership in the United States—and to some extent in Europe—has preferred to have a small handful of authorized voices on Mexico rather than to take the country seriously as a site of cultural and intellectual production. It has been economical and convenient for Americans and others to simply tune in to Carlos Fuentes, Octavio Paz, or Enrique Krauze and to take whatever they say as representative of what Jose María Morelos called "the sentiments of the nation." However, the power to represent Mexico in this way, to embody it in a single intellectual is as dead as the autocratic power of the president.

When he was at the height of his power, president Miguel Alemán wanted the Nobel Peace Prize. President Luis Echeverría tried for the General Secretariat of the UN, and Carlos Salinas wanted to be president of the World Trade Organization. These unkingly desires reflect the nature of presidential power and the limits of presidential biographies: they are not the main axis in the history of Mexico. I like to think that this book is the intellectual counterpart of these desperate presidential moves: the concentration of cultural power in the hands of a few intellectuals has been linked to the authoritarian power of Mexican presidents, and the current democratization and debilitation of the presidential office promises to end this form of "intellectual caudillismo."

REFERENCES

- Krauze, Enrique. 1986. *Por una democracia sin adjetivos*. Mexico: Joaquín Mortiz.
———. 1992. *Textos heréticos*. Mexico: Grijalbo.
———. 1997. *La Presidencia imperial*. Mexico: Tusquets.
Lomnitz, Claudio. 1992. *Exits from the Labyrinth: Culture and Ideology in Mexican National Space*. Berkeley and Los Angeles: University of California Press.
Lynch, John. 1992. *Caudillos in Spanish America, 1800–1850*. New York: Oxford University Press.
Mundy, Barbara E. 1996. *The Mapping of New Spain: Indigenous Cartography and the Maps of the Relaciones Geográficas*. Chicago: University of Chicago Press.
Myers, Jorge. 1995. *Orden y virtud: El Discurso republicano en el régimen rosista*. Buenos Aires: Universidad Nacional de Quilmes.
Pozas, Ricardo. 1993. *La Democracia en blanco: El Movimiento médico en México, 1864–1965*. Mexico: Siglo XXI.
Segovia, Rafael. 1980. "La Decadencia de la democracia." *Razones* 24 (March–April).

Book Reviews

The Creativity of Action. By Hans Joas. Translated by Jeremy Ganes and Paul Keast. Chicago: University of Chicago Press, 1996. Pp. x+336. \$49.00 (cloth); \$19.95 (paper).

Colin Campbell
University of York

This book by Hans Joas, professor of sociology at the Free University of Berlin, was first published in German in 1992 and represents a development of ideas originally expressed in *Pragmatism and Social Theory* (translated by J. Gaines, R. Meyer, and S. Minner [University of Chicago Press, 1993]). Essentially it is an attempt to demonstrate the continuing value and relevance of the American pragmatist strain of social theory (as represented by George Herbert Mead, William James, John Dewey, and Charles Sanders Peirce) by indicating how their ideas can be used to develop a theory of creative action; that is, "a concept of action that is reconstructed in such a way so that . . . [it] . . . is no longer confined to the alternative of a model of rational action versus normatively oriented action, but is able to incorporate the creative dimension of human action into its conceptual structure" (p. 72).

The book itself is in four sections. In the first Joas discusses each major "action theorist" in turn, assessing the individual's approach and noting his mistakes, while also pointing out the extent to which an incipient theory of creative action can be found in his work. Thus Joas identifies Parsons's basic error to be the attempt to emulate an essentially means-end model drawn from economics, while in his discussion of Weber he observes that the concept of charisma can be seen as significant because it highlights the "creative" aspect of human action. In the second section Joas turns to look more closely at social theories that deal explicitly with human creativity, building his discussion around the metaphors of *expression*, *production*, *revolution*, *life*, and *intelligence and reconstruction*, all of which he finds expressed in both the "philosophy of life" theories developed in Europe by the likes of Herder, Marx, and Schopenhauer, as well as by the American pragmatists. Then in the third and clearly the most important section of the book, he outlines the basis of his own theory of creative action. Finally, in the fourth section, he endeavors to demonstrate how this model of action can contribute not merely to the development of a more adequate social theory but also to a resolution of the crisis of postmodernity.

Although there will be those who will wish to challenge the interpretations that Joas offers of the work of the many social theorists that feature

Permission to reprint a book review printed in this section may be obtained only from the author.

in his wide-ranging and erudite discussion, to focus on the issue of the correct interpretation of past and present theorists would be to miss the significance of this book. For, as in the case of Talcott Parsons and *The Structure of Social Action* (McGraw-Hill, 1937), the discussion of other theorists is essentially a preparatory exercise, intended merely to lay the groundwork for his own theory. In this respect Joas is at his persuasive best when it comes to identifying the weaknesses of existing theories of action; although it has to be said that many of these have been noted before. However, his critique is more radical and thoroughgoing than most insofar as he attacks not just the emphasis on rationality but the instrumental means-ends schema as a whole, observing that such a framework excludes not just impulsive and habitual conduct but also all that is autotelic. What is more, Joas rightly draws attention to the fact that the body does not appear in most theories of action, even though "the relationship between personality and that person's body is one of the central issues in the theory of action" (p. 171).

Hans Joas's determination to develop a nonteleological approach to action, especially one that stresses corporeality as well as the situated and biographical character of action, is to be very warmly welcomed indeed. This having been said, there must be some doubt about whether what Joas presents us with really warrants describing as "a theory of action" rather than simply a prolegomenon for one. For, although he identifies the features that such a theory ought to possess, he does not set out their interconnections in anything like the detail and specificity required of a theory. No sociologist, having read this book, would be in any position to go out and apply Joas's theory of creative action. Indeed, nothing reveals the undeveloped state of his "theory" so much as the fact that nowhere are we provided with a definition of action, nor—what presumably might have served almost as well—a definition of creativity. This absence is critical since, given the wide range of theorists that Joas draws upon (many of whom would not normally be thought of as action theorists), the empirical referent for this concept cannot simply be taken as given. Despite this serious omission, this is a very significant work. We can hope that its publication in English will prompt more sociologists to abandon the theoretical cul-de-sac represented by rational action theory (as well as the subjectless normative or so-called social action theories) in favor of more promising lines of theorizing about action. Joas's is certainly one of the more exciting—and creative—of those currently on offer.

Fin de Siècle Social Theory: Relativism, Reduction, and the Problem of Reason. By Jeffrey C. Alexander. London: Verso, 1995. Pp. vii+231.

Michèle Lamont
Princeton University

Fin de Siècle Social Theory is Jeffrey Alexander's response to recent development in sociological (and social) theorizing. In this book, the author tackles everything "post" and improved (including things modern, "anti," and "neo"). At the request of the book review editor, I will only comment on the chapter that has not been previously published and which makes up almost half of the book, the author's *attaque en règle* on the work of Pierre Bourdieu. This is true, genuine, feisty, polemical Alexander, and he takes no prisoners. His argument can be schematized as follows. (1) Bourdieu's critique of structuralism and attempt to restore agency ignores the tradition of practical action theory (namely Peircian pragmatism), with the result that he inflates the originality of his own theoretical contribution while turning structuralism into a "vulgarized enemy-tradition" (p. 135). (2) His concept of habitus "binds actors tightly to the social world" (p. 143) and downplays the relative autonomy of culture (i.e., the existence of universalistic ideals that can be generalized across contexts and exist beyond the hierarchical structures of material life). In a provocative and original comparison between Bourdieu's work and psychoanalytic and developmental theories of the self (p. 145), Alexander argues that in the former, there is no self or moral choice possible since a critical distance from social structure cannot be attained. (3) Bourdieu conceptualizes action as only cognitive and strategic, at the exclusion of communication and collective obligations; affective dispositions and symbolic schemas cannot be motivational sources for action, only "practice as profit-seeking." (4) Similarly, Bourdieu's concept of field is reductionist because it presumes a homology between field dynamic and broader social relations. (5) Bourdieu's empirical work is primarily a means to elaborating discursive commitments. It suffers from the same problems as the theoretical work, namely "the impoverished understanding of meaning, the caricature of motivation, the inability to conceptualize the interplay of sensible self and differentiated institutions in contemporary society" (p. 171). (6) Bourdieu's critical theory is not critical because it leaves no room for differentiating between democratic and totalitarian societies and for capturing the place of moral judgment and deliberation in the public sphere. (7) Bourdieu is less original than he claims because, in fact, he is a closet neo-Marxist who "conceives of actors as motivated by a structure of disposition which merely translates material structures into the subjective domain" (p. 136).

Responding to Alexander's polemical spirit, Bourdieu aficionados will undoubtedly take exception to many of his interpretations, grounding their own reading in the sanctioned guidelines that Bourdieu kindly provides. A fair and dispassionate assessment of the merit of each critical

point would require more space than I have here (and the stakes are high given the intensity of the critique). I will make only a few points. Most important and critical, concerning Bourdieu's closet neo-Marxism, Alexander's critique would be more crucial if he could show that Bourdieu is only a neo-Marxist because our Frenchman's influence is grounded in part in the originality of his synthesis of neo-Marxism, neo-Weberianism, and neo-Durkheimianism (on this point, see Rogers Brubaker, "Rethinking Classical Theory: The Sociological Vision of Pierre Bourdieu," *Theory and Society* 14 [1985]: 745–75). Another weakness is that while the critical synthesis Alexander provides is a useful and provocative thinking tool, he does not pay sufficient heed to available theoretical critiques and should have been more specific concerning the originality (and limits) of his contribution. On a more positive note—and although I should not be judge and party here—I am sympathetic to aspects of Alexander's theoretical and normative critiques of Bourdieu's treatment of moral issues because they are confirmed by key empirical findings of my book, *Money, Morals, and Manners: The Culture of the French and American Upper-Middle Class* [University of Chicago Press, 1992]. Finally, although issues of interpretations will be contested, Alexander's scholarship is generally careful; I could only identify a few instances where it was questionable (e.g., Alexander uses, out of context, Bourdieu's writings on universalism formulated in his epistemological critique of survey research to draw conclusions about his views on universalism in general).

Now that American sociologists have ingested most of what Bourdieu has to offer, we should continue to approach it with less hype, to move away from celebratory modes, and to turn toward more grounded confrontations between theory and empirical terrain. In a cosmic sense, whether Bourdieu is right or wrong is less important than the use posterity makes of his work. In this context, Alexander's essay can be an important stimulus because, just as Bourdieu's writings have done for so many readers over the past 20 years, it helps foreground questions that have remained muted and unseen to this day.

Sociological Reasoning: Towards a Post-Modern Sociology. By Rob Stones. New York: St. Martin's Press, 1996. Pp. xii+257. \$49.95 (cloth); \$18.95 (paper).

William Outhwaite
University of Sussex

Sociological Reasoning is an ambitious and important book. The author confronts the opposition between traditional residually empiricist conceptions of scientific sociology and a "defeatist" postmodernist relativism that stresses invention and style rather than substance. Stones identifies the former image of sociology with a variant of modernism, though he

stresses how far it fell behind the sophisticated insights of artistic and literary modernism. His proposed alternative to the traditional conceptions is based on two premises: a scientific-realist ontology that upholds the reality of the social world and what he calls a "past-modern" recognition of the complex issues raised by the attempt to describe and theorize it. Sociological modernism, in Stones's terms, tended to move rather fast from assuming the reality of the social world to the further assumption that it was just the way the theorist described it—whether in the theory-poor variants of empiricism found in much Anglo-American sociology of the middle decades of this century or theory-rich accounts such as Marxism framed in terms of underlying structures.

Realists in the social sciences have, Stones implies, paid lip service to the complexity of reality and, hence, to the need for our descriptions of it to be similarly complex but then has moved briskly to the implementation of some particular theoretical option, which is given a more or less spurious, spray-on realist gloss. In other words, realism makes definitions truth functional—where for positivism and conventionalism they were matters of arbitrary choice. It then allows a similar free-for-all at the level of *theory* choice on the grounds that philosophical arguments cannot be used to support particular ontologies of the social based on concepts such as class or patriarchy. (Stones suggests that Sylvia Walby's explicitly realist analysis of forms of patriarchy is an example of this use of realism to buttress a particular theoretical option; see her book, *Theorising Patriarchy* [Oxford: Blackwell, 1990]).

What Stones does is to take much more seriously the agonistics of theorizing, the fragility of knowledge claims, and the politics of style and presentation. Perhaps the best illustration of his approach is one of the case studies that make up the final chapter of the book, he rewrites a footnote in Geoff Ingham's book on the City of London (*Capitalism Divided* [Macmillan, 1984]) in a way that makes evident to the reader the necessarily conjectural elements in Ingham's account of Harold Wilson's economic policy.

In the systematic central section of the book, Stones elaborates a phenomenology of styles of theorizing. First, he distinguishes between theories that focus on "agents' conduct," those based on what he calls the "theorist's pattern," and between contextualized and generalizing ("floater") modes of presentation. Thus, Michael Mann's classic account of the development of modern state forms ("Ruling Class Strategies and Citizenship," *Sociology* 21:339–54) is a floater theorist's pattern analysis, where historical detail and specific policies are compressed into a structural narrative (*sjuzet* rather than *fabula*, in the terminology of Russian formalism that Stones introduces in chap. 7).

Chapter 3 outlines a trichotomy between "players," "dreamers," and "despots." Players address specific cases and tend to focus on "actors'" meanings (e.g., Jack Douglas's *Social Meanings of Suicide* [Princeton University Press, 1967]). Dreamers operate to a high level of abstraction (as in Anthony Giddens's *Modernity and Self-Identity* [Polity Press,

1991]), while "despots" (e.g., "modernists") force reality into a tightly formulated bed of assumptions and categorizations.

It is impossible in a short review to bring out the full complexities of these categories as Stones develops them. The upshot, however, is to encourage greater methodological reflection in "the stretch of land between high theory and empirical research" (p. 234). Some readers may find Stones's neologisms irritating; but his stress on what one might call the political stylistics of social theorizing is well taken, and the wealth of sociological and literary examples both illustrate and enliven the texts. My personal quibbles are with Stones's use of the term "sociological modernism" to describe what his own analysis of the sociological tradition shows to include—a broad spectrum of approaches and styles—and a tendency to treat the discipline of sociology (albeit broadly conceived) as what realists call a natural kind (i.e., as having a clearly demarcated identity). There is, however, much food for thought in this splendid book, and I hope it will receive the attention it deserves.

The Social Construction of Nature: A Sociology of Ecological Enlightenment. By Klaus Eder. Thousand Oaks, Calif.: Sage Publications, 1996. Pp. x+243. \$75.00 (cloth); \$26.95 (paper).

Steven Ney
Institut für Technik und Gesellschaft

Why is it that although "green ideas" have been with us for nearly two decades, both industrial and developing countries are still set on an environmentally destructive course? Along with the many economic and political answers available, Klaus Eder, who holds a chair of sociology at Humboldt University in Berlin, offers us a cultural explanation.

Eder's book is a highly theoretical piece of work, which, in its philosophical bent, is quite characteristically German. Ecological reason alone, Eder surmises, may not save us from our present destructive path; after all, ecological reason, the latest form of practical reason, emerged from the same utilitarian discourse that led to environmental pollution. Rather than relying on practical reason, sociology ought to look for the social locus of environmental "irrationality," the modern cultural constructions of nature.

Yet, where do we start? Since theoretical reasoning can do no more than provide "enlightenment on the illusions tied to [practical reason]" (p. 1), Eder suggests that sociology be reconceptualized along the lines of cultural theory. On the one hand, this reconceptualization overcomes the naturalistic bias of classical sociology, which casts nature as something external to society. On the other hand it focuses on how societies symbolically structure their access to nature.

Eder is in line with current trends of extending social scientific analysis with cultural theory. Yet, showing sociology that culture is important is

the academic equivalent of taking coal to Newcastle; there must be something else going on. Indeed there is. Eder also wants to salvage a philosophical idea of rationality (practical reason) in a sociological analysis that makes extensive use of sociocultural constructions of nature. He does this by locating rational social learning processes not in the evolution of rational ideas about nature but in the evolution of cultural constructions of nature. The development of these cultural complexes, whether they are based on magical or scientific ideas about nature, provides people with the means to act reasonably. Thus, Eder has extended sociological analysis: modern ideas about nature can be contrasted with social constructions from other, possibly nonmodern, cultures.

Since everyday cultural practices are key to a critique of ecological rationality, part 2 of the book embarks on a thorough analysis of food taboos and eating rituals. In his review of case studies about food taboos in native South American tribes, Jewish culture, and modern society, Eder draws on work from social anthropologists such as Mary Douglas and Marshall Sahlins. Here, food taboos and eating mores are means of communicating social order. Using this approach, Eder identifies two opposing culinary cultures in modern society: the carnivorous and the vegetarian cultures. These cultures express two fundamentally opposed cultural constructions of nature: the former conceives of nature as an object that serves human purposes whereas the latter sees nature as having intrinsic value. Carnivorous culture, based on sacrificial rituals, is aggressive, dominating, and violent; it objectifies the world in order to manipulate it. The vegetarian culture, in turn, is based on the explicit rejection of blood sacrifice, which Eder tries to trace back to the Greek Pythagorean and to the Jewish traditions; it posits a more cooperative relationship with nature based on the principle of equality. These two cultures, then, provide the basis for the ecological debate raging in the public sphere: the carnivorous culture underpins the dominant polluting faction, and the vegetarian culture strives for a new relationship with nature.

Ultimately, Eder's message is political. He sees the environmental problem as a vehicle for democracy in industrialized states. Environmental ideas have transformed the public discourse on the environment, and we are now in a phase of postenvironmentalism. Environmental frames of reference introduced to the public sphere a decade or so ago have been institutionalized and have become ecological ideology. This transformation brings the two cultures of modernity into conflict; they have to find a new way of coexisting. Eder argues that this has caused a rationality paradigm shift that challenges the legitimacy of the corporatist state: environmentalism has introduced ideas of civic participation into the public sphere.

Klaus Eder's book is an interesting contribution to the environmental debate. Eder devises a rigorous framework in which nature is socially constructed but still enables the analyst to make sensible statements about the rationality of environmental practice. Further, Eder impressively weaves together different strands of the social sciences into a coher-

ent whole thus making a strong case for the theoretical pluralism that is becoming increasingly popular in the policy sciences.

These positive aspects make the final result all the more disappointing. This reviewer cannot help but wonder whether the complexity of environmental thought is ill served with a binary cultural model even if they are, and this is not clear, ideal-types. Moreover, Eder does not always help the reader. The dense prose, the theoretical jargon, and the lack of illustrating examples outside the case studies make for difficult reading. This is compounded by a poor translation from the original German. Nevertheless, Klaus Eder's book is a stimulating piece of work that should serve as a basis for future research in the field.

Marginalized in the Middle. By Alan Wolfe. Chicago: University of Chicago Press, 1996. Pp. xii+291. \$24.95.

Richard Madsen
University of California, San Diego

Alan Wolfe is one of our best social critics. Few contemporary sociologists have been as prolific, as wide ranging, and as capable of reaching a large, general, educated public, and few have offered commentary that is so consistently well informed and well balanced. *Marginalized in the Middle* consists of a generous sampling of the essays he has regularly written for popular intellectual journals, mostly the *New Republic*. But it is more than the sum of its parts; Wolfe has revised and reframed his previously published essays so as to construct a general argument about the role of the social critic in modern American society.

It is less an abstract, impersonal argument, though, than an *apologia pro vita sua*. In the short opening chapter (the only chapter that is completely new), Wolfe describes the "golden age of social criticism"—the 1950s and 1960s. This era is dominated by figures like David Riesman, Philip Rieff, Daniel Bell, and C. Wright Mills, intellectuals whom Wolfe sees as role models not necessarily for their various political positions but for their ability to bridge the boundaries between academic disciplines and to address fundamental national issues with the aid of empirical research. Wolfe thinks that this golden age was founded in "a faith in social science, a belief in democracy, and a strong sense of political commitment" (p. 3). He argues that these pillars of classic social criticism now work at cross purposes. Nonetheless, social criticism is still possible as long as it realistically adapts itself to the changed circumstances of American life.

The subsequent chapters in the book are, essentially, critical essays on the works of contemporary, would-be social critics—prominent members of the academy who are dealing with race, gender, sexuality, immigration, welfare, schooling, public morality, and power. Taken as a whole, these essays show us how Wolfe thinks social criticism should and should not

be done—mostly how it should not be done, because Wolfe tends to focus on authors who, in his view, take extreme, narrow, or one-sided positions. He aims his criticisms at a wide range of targets spanning the contemporary political and theoretical spectrums: for example, feminists like Judith Lorber and Catharine MacKinnon, theorists of race like Andrew Hacker, rational choice theorists like Richard Posner, scholar/policy advocates like Theda Skocpol, and communication theorists like Deborah Tannen. These thinkers have gone at least partially wrong, in Wolfe's view, because they are not sufficiently grounded in empirical data, not based in sufficiently broad theories, and/or have not properly reconciled the tension between describing what *is* and advocating what *ought* to be.

According to Wolfe, good social criticism is "realist" rather than "romantic." Romanticism rejects secular, utilitarian, bourgeois values and celebrates the heroism and fraternity of the poor and the outcast. (Wolfe classifies as romantic not only my own *Habits of the Heart*, but also, for somewhat different reasons than for *Habits*, much of feminist theory and ethnic studies.) "In contrast to romantic critics, realistic social critics try to understand the world around them, no matter how distasteful, politically objectionable, or immoral that world may be" (p. 46). Realism is especially concerned with the life of the middle classes. Wolfe himself aspires to be a realist and advocates that stance for all social critics. Yet, he warns against becoming so realistic that one complacently accepts the world as it is. Realism is a starting point; it humbles the critic with recognition of nuance and complexity. But having been thus humbled, the responsible critic should press on to advocate ways to improve the world.

Exactly how does one strike a balance between *is* and *ought*, between social science and moral rhetoric, between intellectual freedom and social responsibility? Wolfe does not give us clear theoretical guidance here. Assembled from essays written on different occasions in reaction to social critics who have, in Wolfe's view, failed to maintain the proper balance, *Marginalized in the Middle* lacks a strong, unified argument about how to get the balance right. Sometimes what Wolfe says in one chapter undermines what he says in another. For instance, he criticizes romantic nostalgia while building his whole rhetorical structure on the notion that there was once a golden age of social criticism. Although Wolfe sometimes demonstrates exemplary wisdom in his attempt to struggle with the contradictions of contemporary intellectual life, at other times, in my view, he slides into platitudes about the need to have things both ways.

He does, however, cite good examples of good social critics—Christopher Jencks, William Julius Wilson, and above all, David Riesman—who combine an independence of mind, broadness of knowledge, and commitment to careful research. Wolfe was inspired to become a social scientist when he read *The Lonely Crowd* 35 years ago, and he dedicates his book to David Riesman. Like Riesman, Wolfe confounds conventional wisdoms on both the right and left—all the more important in the present time when political boundaries are in such flux—and tries to stand "marginalized in the middle" of a lonely crowd of intellectuals try-

ing to balance a critical individualism with commitment to their national and academic communities.

Genealogies for the Present in Cultural Anthropology. By Bruce M. Knaft. New York: Routledge, 1996. Pp. viii+384. \$59.95 (cloth); \$17.95 (paper).

Michael M. J. Fischer

Massachusetts Institute of Technology

The MacGuffin (a shaggy dog story Alfred Hitchcock used to illustrate that the real story might be elsewhere than the ostensible focus of attention) in this book is, as the author suggests in the preface, reading for the ethnography of Melanesia, particularly that of the south coast of New Guinea. By playing off Melanesia against fashionable "theory" writers, Knaft hopes to enrich both.

In the chapter on Foucault, for instance, Knaft writes, "Since sex and violence are prominent in significant portions of Foucault's historical work and in the ethnography of south coast New Guinea, they provide a substantive as well as an analytic point of linkage" (p. 144). This is a region rich in the diversity of ritualized sexuality (homosexual as well as heterosexual), the use of bodily inscriptions to institutionalize social divisions and relations of domination (nose-, penis-, tongue-bleeding; burning; rubbing with nettles; etc.), head-hunting raids, and "radical and prolonged initiation rites" for children of decapitated adults. Such practices and their cosmological armatures, when read against several generations of ethnographic description and analysis, provide a way to explore the shifting conceptual categories, anxieties, and ideologies of anthropological, missionary, governmental, and political-economic frameworks from 1900 to the present. This list should include most recently the invention of *private* homosexual relations (*mbai*) where missionaries and governments have eradicated *public* ritualized homo- and heterosexual (*pap-isi*) exchange partners. *Mbai*, says Knaft, is "radical resistance to and infringement upon the sexual taboos of the Catholic priests" and illustrates "a subtler change in what Foucault identifies as the way in which the individual as *ethical subject* establishes his or her relation to the rule" (p. 169).

The chapter on postmodernity and Melanesia picks up the ethical in terms of political economy. The ecological and social devastations caused by oil companies and mining and logging operations by big corporations have generated local resistances in many forms, which detailed ethnography can foreground. It might have helped Knaft's case if instead of joining those who choose to read "postmodern-ism" as merely engaging in pastiche, irony, and other aesthetic tactics, he had acknowledged, for instance, Derrida's defense of Marx as prophetic of an intensifying speculation of commodity fetishism (*Specters of Marx*, Routledge [1994]),

which continues to produce the "ten plagues" of unemployment, economic warfare, and so forth, or had acknowledged the generational experiences out of which French postmodern theory arises—decolonization (Algeria), technological modernization (computers), (re)pluralization of French culture (*Le Pen*), and so on. Still, the important point, which Knauft gets right, is to insist that theorists be read against the context of the political, economic, and ethical issues on the ground. Heidegger, thus, fares badly in Knauft's reading, quite apart from his Nazism, for his fundamental asociality and ahistoricity in contrast to Bakhtin.

The chapter on Bakhtin and Gramsci uses the ethnographic examples of changes in the form of sorcery accusations and spirit mediumship among the Gebusi of Papua New Guinea. Comparing these to Michael Taussig's descriptions from South America, Knauft suggests how these accusations and new songs of a particular spirit medium expose and negotiate the social encroachments of capitalist economic relations. This is a rich vein of thought that Knauft uses to foreground the ways in which local agency is developed and contestations of hegemony are pursued. Other ethnographic comparative examples he might have drawn upon include Richard Werbner's account of changing uses of witchcraft in Zimbabwe after the civil war or Anna Tsing's account of sorcery and possession in Malaysian factories.

The chapter on Bourdieu invokes the many ways warfare in Melanesia has been analyzed over the decades: precolonial indigenous collective violence seen as savagery, later as ecologically adaptive or as evidence about the nature of human aggression or warfare in abstract comparative perspective; World War II in Melanesia as generative of cargo cults; postcolonial conflicts as producing coups and gangs. These can be used reflexively to expose the shifting frames of "objective" analysis, Melanesian authors themselves adding their own diverse perspectives. Knauft uses this to criticize what he sees as the weaknesses of Bourdieu and practice theorists: abstraction, failure to deal with agency, or to think through the ways postcolonial constructions of identity "do not boil down to a shared location" (pp. 131–32). Again, a more charitable reading of Bourdieu might be to reinsert his writing into his struggles to come to terms with his historical context—"Algeria 1960" for the essay on the Kabyle house, for instance.

To read Knauft's book for the ethnography can be rewarding, as is the effort to read theory into its historical contexts of production. It seems to me that, were Knauft to invest the same generosity of contextual reading of ethnographic and theoretical work elsewhere as he gives his colleagues in Melanesia, we would have a far more powerful contribution to the ethnographically informed critical humanism he and many of the rest of us have long been calling for. Anthropologists continue to produce a quite rich corpus of ethnography, including regional sensitivity to the changing political economy of late modernity or global capitalism and a sensitivity to contestatory social forces and varieties of mobilizing the diverse threads of culture and power.

Communicative Interaction, Power and the State: A Method. By Frank M. Stark. Toronto: University of Toronto Press, 1996. Pp. xi+227. \$55.00 (cloth); \$19.95 (paper).

James Fernandez
University of Chicago

The classical notion—put forth by the Sophists, rejected by Plato, and ambivalently regarded by Aristotle—that the key to effective governance and a satisfactory integration to community life lies in the powers of persuasion of those who presume to govern and the openness to persuasion of the governed is taken up again by Frank Stark. This is addressed, however, by reference to more recent classics: the work of George Herbert Mead and Kenneth Burke. The focus on these two theorists is complemented by use of Anthony Giddens's notions of power as a property of interaction and by Antonio Gramsci's attention to that balance of coercion and consent characteristic of hegemonic relationships. More recent theory has obscured as much as clarified the arguments of Mead and Burke, and Stark seeks to clear away this accretion and return to the basics of communicative interactionism and the rhetoric of social drama (dramatism). He seeks to do so not in respect to the microanalyses of social dynamics, which are usually understood to be these authors' contribution, but in respect to the much broader institutional contours of public policy formation at the national and international level. The theory advanced here, to be more fully worked out by Stark in work in preparation, is a theory of shared or interactive meaning in national and international community building. It is a theory of how community at these levels is gained and how it is lost. In contrast to theories of *social action* (that of Talcott Parsons, e.g.), the author proposes, in the tradition of symbolic interactionism, a theory of *the dialectic of social interaction*.

Stark's vision is democratic and cooperativistic, and the communicative interaction he has in mind, Meadian in its emphasis on benevolent mutuality or constitutive interaction of self and other without alienation or selfishness, eschews the forcing or seduction of public opinion in the propagandistic mode. He makes an instructive distinction between *propagandistic* regimes, dehumanizing and repressive of citizen reflection, and *persuasive* governments, which encourage reflexive identification of "selves" with the "others" in community. His vision of rhetoric, then, is not a coercive one. "If all rhetoric is a reduction to *perceived* choices, then it is only a kind of psychological coercion." Persuasion is "an enhanced understanding of available choices, in the context of an identification of the rhetorician with the interests of his audience." Propaganda is a denial of choice through a discouragement of the consciousness of choice (p. 29).

The author complements Mead's benevolent vision of community based on a mutual taking of the other with the less muted and more agonistic awareness of faction of Kenneth Burke—a vision played out in Burke's dramatism. This perspective, dramatic and hence agonistic by

definition, constitutes a method, which is to say a template or "intelligible frame," for parsing, interpreting, and narrating the scenarios of community interaction in which there is an ongoing struggle to define reality. Central to this perspective is Burke's well-known dramatic pentad (later hexad) of *act, scene, agent, agency, purpose* (and, later) *attitude*. This is a method, template, or frame capable of unraveling and bringing to satisfactory explanation the complex communicative interactions that occur in institution building.

Stark makes a good deal of the choice of basic organizing metaphors facing any student of social interaction and enumerates the advantages of the dramatic trope over the biological and mechanical models current in a historically oriented social science on the one hand and in functionalist inquiry on the other: (1) It is a trope that avoids "path dependency" or "linearity" of understanding and instead focuses on reflexive activity; (2) it is a method for relating various philosophical views of social action to each other; and (3) it is a means for studying rhetoric in communicative interaction. In short, it is a trope that uses that which man himself has created, drama, to understand broader acts of institutional creation. It will be recalled that Vico argued reflexively in the *New Science* that man is only able to fully understand that which he himself has created. Stark also makes a distinction between the "dramatism of the act" of Mead and Burke and the *dramaturgic* approaches of Goffman and others which downplay reciprocity and mutuality and overemphasize the manipulative and the secretly coercive in communicative interaction.

Using the dramatic model, Stark analyzes the scenarios of three attempts at national and international institution building (attempts at federation) in Canada and Cameroon. He argues that the method of dramatism complimented with fieldwork carries inquiry more deeply into the negotiation of universes of discourse and the complex relation between rhetoric and institutional results. Like any metaphor, the metaphor of dramatic scenario for social life risks concealing as much as it reveals—the material interests at play, for example. But Stark's protagonism goes a long way toward lifting the reader's disbelief and persuading him of the dramatic point of view.

Rational Choice Theory and Religion: Summary and Assessment. Edited by Lawrence A. Young. New York: Routledge, 1997. Pp. xiii+187. \$59.95 (cloth); \$17.95 (paper).

John R. Hall
University of California, Davis

Religions often traffic in miracles, the supernatural, matters of faith, and revelation. Sociological ideas about religion—from Durkheim and Weber to the secularization theory that posits an opposition between religion and

modernizing rationality—treat religions as paragons of culture formed by communities that encompass individuals in webs of meaning transcendent to their worldly motivations. On any of these grounds, religion represents an acid test for rational-choice explanations, which focus on the decisions individual actors take in order to maximize returns on their values based on the best information they can summon at the time.

It is, thus, of no small significance that since the late 1970s a group of social scientists led by Rodney Stark has elaborated an increasingly sophisticated rational-choice theory of religion. Lawrence Young's edited volume, *Rational Choice Theory and Religion*, brings together 10 papers from a 1994 conference in Sundance, Utah, which offer retrospective accounts of the theory's development and prospective assessments of its potential. This book is tremendously important for the sociology of religion and as a harbinger of rational-choice theory's wider sociological prospects.

The book's first four chapters encapsulate central debates in rational-choice theorizing about religion. Rodney Stark, with William Sims Bainbridge, begins with an engaging saga of his efforts to develop a propositional testable theory. Their central idea is that although some rewards are not available in this world, religions can offer "compensators" that "provide an explanation about how the desired reward (or an equivalent alternative) actually can be obtained, but propose a method for attaining the reward that is rather elaborate and lengthy" (p. 7). The next crucial step is to view religious organizations as firms and seekers as consumers within "religious economies" that tend toward a free market or monopoly. Rodney Stark, William Sims Bainbridge, Larry Iannaccone, Roger Finke, and others have used such formulations to explain the failure of secularization theory, to account for the success of strict religious organizations, to model church-sect dynamics, and to explore "market regulation," market niches, firm efficiencies, and religious competition—especially in "the churching of America."

After the opening essays, chapters by other sociologists of religion and social theorists all are careful to praise the rational-choice Caesar, but they also ponder symptoms of malady. They propose infusions, amputations, organ grafts, and sex change operations. In the end we wonder whether we know him anymore: Is he alive or dead, ready for burial, or already reincarnated into an alternative life form?

To begin with, there are questions internal to rational-choice theorizing about religion. Iannaccone and Finke both assume preferences are stable from person to person and over time. But Michael Hechter calls this assumption "overly restrictive" (p. 152) commenting that "rational choice theorists have gotten themselves into a rut by treating every group as if it were a firm, and by focusing their models exclusively on the attainment of private, instrumental goods like wealth" (p. 157). Randall Collins draws the religious-compensators assumption into question, arguing that religions offer real social rewards (p. 170). Other contributors raise numerous empirical challenges to rational-choice expectations. Stephen

Warner questions the theorized relationship between strictness and organizational strength, Nancy Ammerman challenges rational-choice predictions about differential religious participation of the wealthy and the poor, and Mary Jo Neitz and Peter Mueser suggest that European religious monopolies may be partly driven by demand and not simply by enforced market closure.

Intriguing questions also emerge about the relation of rational-choice to other theories. Stark introduces socialization, Iannaccone brings in religious human capital, and Derek Sherkat wants to incorporate seduction, sympathy, and the social embeddedness of religious economies. Then, several contributors argue that the natural science cause-effect model fails to identify mechanisms because it refuses the thick description made possible by ethnographic methods. Warner suggests that the exciting new paradigm is centered not on rational-choice theory but on the distinctive flexibility of American religious disestablishment.

Finally, several contributors are struck by how much is left "outside of the analytic frame," as Neitz and Mueser put it (p. 110), when religion is viewed through the rational-choice lens. As they observe, gender is virtually absent from the rational-choice analysis, and the very notion of rationality could benefit from elaboration via Alfred Schutz and Harold Garfinkle. Further, what Stark glosses as "explanations of the human condition" (p. 13) and what Ammerman calls "content" (p. 125)—the meanings of different kinds of religiosity—are strikingly rare in rational-choice discussions. With Lawrence Young's invocation of the phenomenology of religion and Collins's discussion of Weber, it becomes increasingly evident that rational-choice theory has difficulties theorizing culture. As Young and Collins both argue, such analysis is not inherently incompatible with rational-choice theory. But I expect synthesis would yield something other than rational-choice theory as presently construed—an understanding of culturally nuanced "rationales" meaningful for participants in distinctive religious communities.

Rational-choice theory is so resilient that it already has broken its own bounds. The great promise is that it will improve debate about diverse explanations that can be formulated within its discursive matrix. The alternative—pursuing a narrow economic model of rationality—can too easily become an ideological glorification of how the free market's hidden hand produces religion that must be good, since it is consumed. This Panglossian vision mistakenly slips from empirical analysis of choices to a normative embrace of one institutionalization of religiosity over others—a shift for which rational-choice theorists lack theological or charismatic qualification.

Congregation and Community. By Nancy Tatom Ammerman. New Brunswick, N.J.: Rutgers University Press, 1997. Pp. xvi+434. \$55.00 (cloth); \$24.00 (paper).

Rhys H. Williams
Southern Illinois University, Carbondale

Nancy Ammerman's latest work does not lack for ambition. Weaving together conceptual strands from the sociologies of community change, organizational ecology, religion, and culture with data from a comparative study of 23 congregations in nine communities, Ammerman explores how changes in the surrounding community affect religious congregations. In particular, she is interested in how and why some congregations adapt while others do not. The data include ethnographic, survey, interview, and archival material collected by Ammerman and a team of 16 researchers.

Ammerman handles this much material through a number of organizing frameworks. There is regional variation; more important, the communities were chosen because they have recent experience with economic, cultural, or social structural transitions. The subject congregations were selected to represent a variety of traditions (all Christian), governance politics, sizes, and demographic makeup. Finally, the chapters on the congregations themselves are organized by the response to change (persistence and decline, relocation in place or identity, or three kinds of adaptation), while each congregation is discussed in terms of its "resources" (buildings, finances, and human capital), "structures of authority" (both external denominational and within the congregation), and "culture" (generally, ideas and practices that represent congregational identity).

Perhaps inevitably, the book has a sprawling feel to it; it is difficult to summarize. If there is one sociological "rule" that emerges, it may well be that those congregations that adapt to change—and thus seem to survive and even prosper—must deal with the challenges openly. This involves conflict over program, leadership, and identity. But it is only by engaging such conflict directly and working through it that congregations can navigate their changing environments. It is a lesson useful for religious congregations, secular voluntary associations, and social movement organizations.

Other important insights include the necessary but insufficient character of material resources such as buildings and budgets. Without the will and skill to "mobilize" such resources, they can be obstacles, as constraining as they are enabling. Ammerman also shows that the connections between various elements of organizational culture are contingent upon ideological and identity work. Conservative theology may point congregations toward particular types of programs, members, or politics, but no avenue is determined or invariant. And the importance of the physical environment, from the character of church architecture to the attachment to a specific location, is a corrective to "place-less" sociologies.

Thus, sociologists of urban life or of organizations who ignore religious congregations will learn much from this book. Scholars of macro social change should read this to watch the process at the ground level. But there is also much to caution those who are too optimistic about religion's place in American society or its role in transforming modern life—the prime mover in this book is the marketplace, with congregations either swimming with the current or stubbornly facing upstream. In that regard, the book's general tone accents the upbeat; even the chapters are arranged in an order of ascending "success." On one hand, this is justified as Ammerman has found a number of success stories that display an impressive array of innovative action by a wide spectrum of ordinary church-going people. And as she notes, congregations continue to be integral to the formation of social capital, community networks, service provision, and the processes of civic engagement.

On the other hand, the most vexing problems in contemporary America still haunt these congregations. The racial divide and, to a lesser extent, the differences represented by new immigration are generally reproduced here. All the congregations struggle with some kind of integration; those who confront racial integration openly have done so after changes in the community forced the issue and even so win only minor victories. The demands of the economy on people's lives, from draining work schedules to dislocations to alienation, are also beyond the reach of the congregations' efforts at social change. Not surprisingly, those congregations adapting most effectively to the changes in their communities are generally (though not completely) those congregations with large numbers of conventionally "successful" members: college educated, middle class, often white, with organizational skills.

Ammerman's final chapter draws together many public issues and their attendant literatures. Her critique of "modernization" theory with its implied secularization is not particularly new, although it has power at the end of this impressive study. Ammerman then follows Nancy Fraser, Patricia Hill Collins, and others in noting how modernity's supposedly rational "universalism" actually marginalized subaltern groups. She argues instead for an associational pluralism (my term) where a certain amount of voluntary sectarian "particularity" allows people to develop the collective identity and resources needed to participate fully in the larger society.

It is a point with which I am sympathetic, but I find it a bit too easy to contrast the failures of modernity's reality with the potential advantages of "models of particularity" that "need not" mean division and alienation (p. 359). This is true, but could not one argue that modernity's universalism "need not" subjugate marginal groups? Much of the particularity I found in the book effectively kept "Euro-Americans" away from African, Asian, and Latino Americans. Such a stance makes it difficult to discuss, as Ammerman does, "community" without specifying whose community or to claim that congregations have "legitimacy" without noting to whom.

These are, of course, dilemmas for all of us interested in a democratic left of both voluntarism and inclusion—and Ammerman offers us an exemplary sociology of organizational and community change that reveals both the potential and the pitfalls.

Religion in the Age of Decline: Organization and Experience in Industrial Yorkshire, 1870–1920. By S. J. D. Green. New York: Cambridge University Press, 1996. Pp. xv+426. \$69.95.

Steve Bruce
University of Aberdeen

Historians and sociologists of religion usually talk past each other. Historians distrust the sociological fondness for grand theory, and sociologists accuse tree-loving historians of failing to recognize the forest. Green is a historian who understands sociology and who has combined the best of both disciplines to produce what is simultaneously a very detailed local history of church life in Edwardian West Yorkshire and an important contribution to the bigger issue of the extent and causes of the decline of religion in modern Britain.

Halifax, Keighley, and Denholme were transformed in the 19th century by the industrialization of the clothing industry. In the period of Britain's imperial heyday, the nonconformist denominations (the source of the most detailed records and, hence, the main focus of the study) grew with the population; they decorated the streets with their buildings and proliferated debating societies, Sunday schools, lending libraries, sports clubs, sewing societies, and the like as they sought to enroll local populations in the associational ideal of the church as an inclusive and all-encompassing community.

What we would now call "mission-drift" probably backfired. The benefits of any of these extensions of church life for evangelism were negligible, and the advantages of institutional proliferation in deepening the commitment of one generation were lost in this century when the growth of attractive secular leisure possibilities gave people a choice between secular and church-based provision. Choosing the secular over the religious option in one part of their lives became a habit. The churches declined.

Many fascinating issues are raised in the details of a period that began with exaggerated optimism about Christianizing Britain and ended with widespread (and fully justified) pessimism about retaining any social influence. For example, Green painstakingly documents the changes in methods of funding the religious life. As churches tried to cope with their deficits (many produced by overambitious building programs), each innovation seems to have served mainly to undermine the previous mode of fund-raising so that the total revenues remained much the same. However, the innovations were arguably worse than useless because the shift

toward raising money from circles outside the church membership through the sale of goods at bazaars and the like broke the moral link between church involvement and church funding.

Green's study has important implications for a number of contemporary debates. First, it offers a firm riposte to those revisionists presently engaged in trying to argue that there has been no major change in the popularity of religion in Britain and that secularization is a sociological myth. As the title boldly states, this is a study of decline. Second, it makes a particular point about the links between supply and demand. Against the expectations of the rational choice theorists, an increase in the number and variety of outlets did not produce an increase in demand. Robin Gill (*The Myth of the Empty Church* [SPCK, 1993]) has drawn on equally detailed studies of church attendance and church building to argue that overprovision actually hastened decline by creating half-empty buildings (which were themselves a discouragement to potential attenders) and by saddling congregations with vast debts. Green's study confirms the second point. The impact of diversification also runs counter to the expectations of the rational choice theorists. Far from multiplying reasons to remain affiliated with the churches, diversification reduced their distinctiveness and forced them to compete on increasingly unequal terms with the works of the devil.

As befits a historian, Green is judicious with his conclusions. I would want to draw from his rich account of what seems a fairly typical part of industrial Britain the broader conclusion that, even if they work for the United States, supply-side explanations of religious behavior patently fail for Britain. Mainstream Christianity did not decline in Britain because the producers failed to innovate, were deaf to the desires of their potential consumers, or were cushioned against shifts in the market by state-provided revenue. It declined because there was insufficient demand and because the innovations designed to stimulate such demand were counterproductive. In one or two places Green's detail threatens to overwhelm the reader, but if sociology's grand theories are to survive the criticism of showing disdain for the evidence, then it is this sort of case study that we must encompass.

Glorious Battle: The Cultural Politics of Victorian Anglo-Catholicism.
By John Shelton Reed. Nashville: Vanderbilt University Press, 1996.
Pp. xxiv+357. \$34.95.

Gene Burns
Michigan State University

I finished *Glorious Battle* convinced that the Church of England is a much more interesting institution than I had realized, given that to some extent it continues an internalized struggle over the legitimacy of the Reformation. Reed's book demonstrates that a central controversy between

the 1840s and 1890s was whether the Church of England was primarily Protestant or Catholic. In the Anglican context, the term "Catholic" implied not fealty to Rome, but a universal Christianity that preceded the Reformation but was not necessarily closely tied to the Vatican. The Anglo-Catholics were an Anglican Church faction who believed that, through carelessness or excess, Anglican belief and ritual had come to accept post-Reformation practices and corruptions that had nothing to do with the true tradition of English Christianity. Anglo-Catholics were particularly prone to conversion to Roman Catholicism, but the central tenor of the movement was to view Roman Catholicism as a sister church, not a mother church. (Opponents of Anglo-Catholicism, on the other hand, focused on the danger of "popery.")

Intellectually, Anglo-Catholicism originated among Oxford University intellectuals of the 1830s and 1840s, most prominently in the person of John Henry Newman. But Reed focuses on the widespread, practical movement for changes in everyday Anglican Church ritual, a movement that outlived the original Oxford debates.

Anglo-Catholic intellectuals emphasized doctrine such as baptism's elimination of original sin and the real presence of Christ in the Eucharist. Such doctrine, in their view, in turn necessarily implied certain specific sacramental and symbolic observances. But it was the ritual itself that caused the greatest controversy, whether or not partisans were aware of doctrinal justifications. And so at times enormous division surrounded such matters as whether a minister wore a surplice or a black gown while preaching, which direction he faced during Communion, and whether candles were lit during services.

For outsiders, it is initially puzzling why what seemed to be religious hairsplitting could be the occasion of riots and heated Parliamentary debates. But after focusing on the specific ritual practices during the first few chapters, Reed provides a compelling analysis of the social and political context that made such changes so divisive. For example, the dramatic growth of urban areas required the building of new churches, which, in densely populated areas, gave Anglican laity a choice among several nearby churches. Ministers in effect became competitors for parishioners, a situation encouraging diversity and innovation, including Anglo-Catholicism, which began as a primarily urban movement. Also, battles over Anglo-Catholicism must be understood partly in the context of the changing boundary separating church and secular authority; Anglo-Catholics benefited from the fact that there was little consensus on the location of that boundary. Thus, while secular courts in theory ultimately could settle disputes within the established church, in fact they hesitated to do so. Within such a context, there was no ultimate authority both willing and able to enforce church discipline. Disraeli did at one point sponsor a law to suppress Anglo-Catholic ritual, but it dramatically backfired: when some Anglo-Catholic clergy were imprisoned based on their conduct of church services, their martyrdom helped draw further support for the movement.

Perhaps the book's most interesting discussion concerns Anglo-Catholicism's challenges to Victorian sex roles. Reed presents evidence that women were particularly attracted to the movement, whether as participants in the revival of religious communities of nuns or as parishioners. The practice of solitary confession before a priest particularly scandalized patriarchal opponents of Anglo-Catholicism: it meant that women might discuss family matters without the presence or permission of father or husband. And, while Reed appropriately notes that our current conceptions of homosexual identity did not exist in Victorian England, it is striking that opponents commonly associated Anglo-Catholic ritual, and male Anglo-Catholics, with an effeminate orientation. Such opponents—troubled, for example, by Anglo-Catholics' valuation of priestly celibacy—hurled charges that we, today, would consider homophobic.

Overall the book is an impressive blend of social history and the sociological study of political culture. I would have just two quibbles. First, most readers will find the first few chapters confusing. They ostensibly provide an overview but take the form of a somewhat disjointed examination of different aspects of Anglo-Catholicism, without explaining, or providing much context for, many of the terms, events, or political groupings discussed. Most such points, however, are ultimately clarified in later chapters or in the indispensable glossary. Second, while the book includes rich sociological discussion of many facets of Anglo-Catholicism, it is difficult to say what are its *central* sociological theses. Still, these are indeed quibbles: most important, Reed's deep empirical analysis of the political, cultural, and even sexual context of religious conflict is thoroughly researched and provides valuable comparative data for sociologists and historians.

Radical Religion in America: Millenarian Movements from the Far Right to the Children of Noah. By Jeffrey Kaplan. Syracuse, N.Y.: Syracuse University Press, 1997. Pp. xviii+245. \$39.95 (cloth); \$16.95 (paper).

Frank Lechner
Emory University

Apocalyptic millenarianism, the dark side of the American belief in progress, has drawn renewed public attention due to a recent series of violent incidents apparently involving people who hold this worldview. Kaplan thinks the attention has focused unfairly on the violence associated with groups preparing for the imminent end of times, and he intends to set the record straight by presenting a more nuanced picture of several fringe groups. His research was specifically inspired by a desire to correct what he calls the Anti-Defamation League's "manichaeian construction of Christian Identity" (p. 176). His argument in this revised dissertation is

that groups that believe the apocalypse requires assistance from believers are not usually disposed to engage in violence. They do share a common "cultic milieu," though incessant disputes fracture the potential "confederacy of seekers."

Kaplan supports these points with a review of published sources and interviews with leaders from three groups: Christian Identity (based on revisionist Christian eschatology with distinctive racial overtones), Odinism/Ásatrú (based on a reconstruction of a Nordic-Germanic faith), and B'nai Noah (based on gentile adherence to the seven Noahide commandments). In each case, he notes that the membership is small and ever changing, that the "theology" is subject to continual revision and that leaders compete for influence via frequent arguments. Most people involved in these groups, he suggests, are "quietist" rather than "activist," even though their rhetoric occasionally comes across as violent. His message is that, extremist members aside, these groups constitute no clear and present danger to mainstream society. Even those who aspire to transform society generally refrain from direct action, since millenarians are "canny judges of the prevailing balance of power" (p. 171) who are not sufficiently isolated to become violently radical.

Why, then, are these groups significant, or at least thought to be significant, by the public and by government officials? In a separate chapter, Kaplan attributes the mistaken perception of the "dominant culture" to the success of the anticult movement, most notably the Anti-Defamation League (ADL) in presenting a caricature of a racist religious fringe. Kaplan does not hide his animus against the anticult movement, which he disparagingly and repeatedly describes as the "high priest of an esoteric god" (p. 129). His pervasive anti-anticult position is somewhat surprising. Kaplan's own evidence shows significant racist and antisemitic tendencies in at least two groups, which would seem to merit the ADL's critical monitoring. Of the groups he studied only Christian Identity has attracted sustained attention, so that Kaplan greatly overstates the anticult movement's role in what he regards as the demonizing of fringe groups.

To most sociologists, this book will be disappointing. Without much narrative verve, Kaplan mainly delineates variations in belief and recounts events in the recent history of the three groups. He does not systematically draw on the sociological literature on social or new religious movements. Apart from scattered comments, he does not tell us how members are recruited (if that notion even applies here), where groups are most active (there is no actual mapping of the groups or their networks), or what accounts for ideological shifts. At various points, for example, where he suggests that Christian Identity eschatology matches the everyday lives of its adherents (p. 50), he misses an opportunity to provide much-needed new empirical evidence. Even though one of his purposes turns out to be to criticize the anticult movement, he barely addresses the ongoing scholarly debate on that subject. Kaplan also says little about

how his distinct anti-anticultist point of view may have influenced his selection and interpretation of evidence. For that reason alone, his account deserves skeptical treatment.

In short, the quality of this work does not match the intrinsic interest of its subject. The book may serve sociologists as a potentially useful source rather than as an authoritative guide. The topic merits a more dispassionate and analytically probing study.

Transnational Religion and Fading States. Edited by Susanne Hoeber Rudolph and James Piscatori. Boulder, Colo.: Westview Press, 1996. Pp. viii+280. \$67.50 (cloth); \$20.95 (paper).

David A. Smilde
University of Chicago

In the introduction and conclusion to this edited volume, Susanne Hoeber Rudolph argues that transnational religious networks have become important forms of transnational civil society. Across the boundaries of the nation-state, transnational religious networks create solidarities, meet individuals' and communities' needs for socioeconomic reproduction and a meaningful universe, and provide "a space for self-conscious, organized actors to assert themselves for and against state policies, actions, and processes" (p. 11). As she sets it up, the validity of her argument depends upon the truth of two basic propositions: first, that civil society is a theoretical construct useful for theorizing the relationship between the state and religious associations both in Western and non-Western contexts and, second, that it is so for transnational religious networks as well. Do the volume's contributors support these propositions?

For the most part, the chapters on Islam indeed reveal, in non-Western contexts, religious institutions with relative autonomy vis-à-vis the state. Don Baker's discussion of East Asia, on the other hand, shows that religion as a separate sphere whose autonomy the political community must respect is not part of the cultural tradition of Korea, Japan, and China. Over the past several centuries, the political powers of East Asian countries have granted autonomy only to those religious associations with international connections because doing so is essential to maintaining esteem in the international community. Thus, while Baker casts doubt upon the universality of "religion as civil society," he shows with unusual clarity the way in which transnational religious networks can generate it where it previously did not exist.

What role do transnational networks play, however, in contexts in which religious formations already have a degree of autonomy vis-à-vis the state? Here the chapters on Islam are revealing. Eickelman argues that while Islam is an important element of civil society *within* states, transnational connections are often superficial. International events and meetings often serve more to shore up the prestige of religious elites

among their local followings than to establish real international networks, and frequently, what seems like international cooperation is actually emulation and the spread of a transnational message. Similarly, Kane shows that while the West African *turuq* are a form of civil society within their own nations, their transnational activity is mainly restricted to commerce and promotion of vague notions of transnational Islamic identity. In both of these portrayals, and in Cary Fraser's chapter on Saudi statecraft, the single most important form of transnational relationship maintained by individual Islamic groups is support from foreign states. Thus, if the individual chapters do not reveal a neat correspondence between civil society, religious associations, and transnational networks, they richly illustrate the importance these networks have in the multiple possibilities of the relationship between states and religious associations.

The most problematic aspect of the volume is the historical narrative Rudolph uses to frame it. In this narrative, nation-states came to monopolize social and cultural domination during the 19th and 20th centuries, while, in the contemporary postmodern era, the reduction of the state is accompanied by an increase in the importance of other forms of domination. Several contributors cast doubt on the first part of this narrative. Fraser shows transnational Islam to have been a part of Saudi statecraft from the beginning of the regime in the early 20th century. And Don Baker speaks of the transnationalization of religion and resulting increase in autonomy as running parallel to East Asian state building over "the last couple of centuries."

The second part of the narrative, on the other hand, relies on a zero-sum notion that should be doubted. (While Rudolph openly criticizes zero-sum metaphors, she really seems to be against "winner-take-all" metaphors: "The process being described here is not the collapse or demise of states but rather the thinning of their effect, function, and finality" [p. 12]). The idea that the contemporary efflorescence of transnational religion has come hand in hand with the thinning of a state that once was the locus of agreed upon norms is true for the industrialized world and some "less developed" countries. But in many areas of Asia, Africa, and Latin America, the nation-state was never more than one more competitor for people's loyalty and obedience. In Latin America, for example, while Rudolph's narrative would certainly fit, say, Mexico, it just as certainly would *not* fit, say, Guatemala. In the latter case one could even argue that the development of transnational evangelical Protestantism has paralleled, even contributed to, the unsteady consolidation of the nation-state.

Despite these reservations, *Transnational Religion and Fading States* is, in my opinion, what an edited volume should be. Rudolph's statements are erudite, carefully reasoned, and agenda setting. And all of the individual chapters are by first-rate scholars who give original empirical treatments of the issues presented by Rudolph. The collection has an exciting potential to stimulate those who work on civil society and religion to look at its relationship to globalization and to stimulate those who work on

the globalization of religion to analyze how it affects the concrete issues of sociopolitical life.

The New Temperance: The American Obsession with Sin and Vice. By David Wagner. Boulder, Colo.: Westview Press, 1997. Pp. x+226. \$60.00 (cloth); \$16.00 (paper).

John C. Burnham
Ohio State University

David Wagner has written a tract to try to counter a current temperance consensus that exists among people of both the conventional Left and the conventional Right. They believe, says Wagner, that governmental powers should interdict and discourage certain behaviors, behaviors that from some points of view were and are private: particularly substance use/abuse, some sexual actions, food and fitness practices, and speech that is not "correct." Wagner correctly points out that temperance crusades do not involve just the substance or behavior that advocates of moderation or prohibition target. The crusaders always exaggerate the dangers and tie the sanctioned materials or actions to other substances and behaviors. And they share Wagner's basic insight that vices and incorrectness have always hung together in social practice.

But if one demystifies temperance crusades, as Wagner attempts to do, what are the reductionistic terms to use on present and past condemnations of substance use, sexual behavior, and other associated "problems" such as violence? Wagner employs Foucauldian constructions along with bright social analyses from a number of recent writers, especially Barbara Ehrenreich. Although Wagner is on the faculty of social work and sociology at the University of Southern Maine, his book is not a technical work in sociology but rather a polemic advocating a particular point of view: attempts to regulate more than a very minimum of behavior are oppressive.

Wagner expresses distress that by the 1980s, both the Right and Left were collaborating in censuring such behaviors as using sanctioned drugs, "having sex" with whomever, whenever, and however, or even eating junk food. He portrays the Right in conventional terms but adds in the idea that bourgeois middle-class standards exploit lower classes in terms of status as well as money. Wagner takes particular pains to show how the 1960s Left was co-opted by temperance crusaders: "Even when liberals and feminists held to their original ground (e.g., with respect to abortion and gay rights), they did so only by finding narrower (nonsexual and nonlibertarian) rights through which to defend these issues" (p. 135).

Where Wagner's work is particularly interesting is in his contention that the economic crises of 1973 changed the social configuration of the United States so that middle-class people of both Right and Left reacted to threats to their economic status by blaming the usual victims, the un-

derclass who were the source of alcoholism, violence, teen pregnancy, drugs, homosexuality, and then AIDS—as well as other unhealthy habits. Many feminists denounced pornography, and gay politics made flaming drag queens into a middle-class minority group—while everyone denounced drug dealers, child abusers, violence on television, and other parts of a familiar complex of behaviors. In short, Wagner utilizes conventional labeling and deviance theory to explain events of the late 20th century.

Against the temperance crusaders, past and present, Wagner sets up a lost ideal of pleasure seeking for its own sake. The ideal is taken from his personal memories or constructions of the 1960s. Unfortunately, his advocacy of pleasure as a justifiable goal is expressed in stereotyped terms. Wagner believes that people who respond to advertising are exercising agency, and he is afraid of government (and possibly ecclesiastical) inhibitions on that agency, not commercial influences.

Despite many passages with sound and even arresting argument, the book will have a much diminished effect because of serious flaws. Neither the author nor the publisher obtained good peer reviewing—or, one suspects, any at all. As a result, the book is riddled with errors and gratuitous omissions of all kinds: not only grammatical, spelling, redundancy, and other copyediting gaffes, but large numbers of totally needless factual errors and distortions. The incidence is so high as not to be a minor consideration. Not only sociologists but the public health personnel who have led much of the new temperance will fault the treatment of epidemiological statistics. The whole area of communications in particular would have helped the book greatly, and in general Wagner is best in those portions in which he cites sociological literature.

Perhaps *The New Temperance* is best viewed as a journalistic argument for a vision of the 1960s that the author senses is heading for extinction. Wagner is essentially and mostly anti-institutional in his orientation. His goal is to overturn the social system. That is very different from overturning the social hierarchy. It could even be argued that Wagner is a dupe of those who profit from the consumer culture. The sex revolution, for example, as historians have pointed out, was never radical because it did not threaten class structure. Instead, fussing about who did what and with what and to whom distracted some troublemakers into thinking that they were really Left radicals. Neither the sexual libertine nor the friendly drug dealer nor denunciations of abstract “power” in fact threatened the everyday practice of economic and social privilege. Nor will this book, although Wagner does ask reformers to focus on social justice instead of the work ethic.

Echoes of the Past, Epics of Dissent: A South Korean Social Movement.
By Nancy Abelmann. Berkeley and Los Angeles: University of California Press, 1996. Pp. xviii+306. \$48.00 (cloth); \$18.00 (paper).

Hyun Ok Park
University of Michigan

What is modernity in Korea? Interpreting discourses and practices of the Koch'ang movement, Abelmann explores "the contours of the debate on modernity and its touchstones—democracy and capitalism" (p. 113). The movement developed in the mid-1980s, when the farmers in Koch'ang of North Ch'olla province demanded the distribution of land that they had cultivated for the Seoul-based Samyang corporation since the Japanese colonial period. For Abelmann, the epics of the movement were shaped by how the farmers remembered the colonial past and the national division and how they understood democracy and economic development under authoritarian rule. Her study differs from two prevalent paradigms used to study Korea. Development studies has skewed the inquiry of modernity into an analysis of the post-1960 economic development attributed to the state's draconian leadership. The postcolonial paradigm, a rather recent perspective in the study of Korea, ascribes the strong state as well as conglomerate-centered economy to the Japanese colonial reforms. Breaking with both of these perspectives, Abelmann concentrates on the unruly practices of people rather than the rational choice of political and economic institutions, on the discursive contours of capitalism rather than take-off and cycles of capitalism, and on an imagined "genealogical relation to the past" (p. 26) rather than an ahistorical narrative.

Counterhegemonic legacy, for Abelmann, lies in "the conceptions of minjung [people]—culture and practice" (p. 113). She argues that "minjung was not merely a stratum in a divided society or a legacy of oppositional activity" (p. 25), as many theorists postulate, but a cultural and historical subject that "imagined a horizontal community or an indigenous cultural socialism" (p. 25). Their quotidian resistance was an attempt to reconstruct their own cultures and national community, not just a struggle to recover national independence or overturn political power. *Minjung* expresses in cultural forms fragmented, real, and mythical relations between the present and the past back to the late-19th-century Tonghak Peasant Revolution and other anticolonial resistances.

Abelmann defines the Koch'ang farmers movement as a *minjung* movement. If two complementary approaches do expound the complexity of a *minjung* movement, one is to compare the Koch'ang movement with other popular movements, including those of the urban poor and workers. The other is to venture on a case study, which is Abelmann's strategy. Chapters 6 and 8 offer an excellent ethnography of how farmers and intellectual activists exchanged dialogues, unlearned official history, and transformed ideological images of each other, even while they kept un-

bridgeable tensions between them. Despite her emphasis on discursivity, she delves into social structure and its relations in movement participation when she reveals that while "wealth correlated inversely to movement participation" (p. 179), the Koch'ang movement cut across some material and political interests and kinship ties with Samyang. Abelmann's attention to gender offers fascinating insights. By sending their wives to the protests, some village members saved face both for the movement and its opponents. In other words, the farmers' own negotiations mattered as much as the discursive and structural conditions, a contribution to the literature of new social movements.

Abelmann's book breaks new ground for the analysis of *minjung* and South Korean social movements. I do, however, have a few critical observations. First, the relationship among subjectivity, discursivity, and imagined memory needs more clarity. While her emphasis lies in discursivity and imagined experience, subjectivity is also figured by institutional arrangements and real historical experience. If *minjung* is a historical subject with a history of passivity and weakness as Abelmann writes, then more discussion on how the Koch'ang farmers in the institutional contexts of the mid-1980s reconciled their protests with their previous silence would provide a less linear concept of *minjung* consciousness, a characteristic of *minjung* literature and the Koch'ang farmers' imagined subjectivity.

Second, the historicity of concepts warrants reflection, when Abelmann characterizes the Koch'ang movement not as tenancy but as a citizenship movement. In articulating a more inclusive vision, the farmers faulted the official modernity that demanded sacrifice for nation. The term *citizenship* is a universal concept that transcends a nation-state's distinctive relationship among individuals, society, and the state, whereas *minjung* with its own temporal and spatial roots signifies collectivity rather than individuality. More discussion on these would enhance our understanding of citizenship and the formation of self, a rather unexplored dimension of modernity in her book.

This conceptual complexity notwithstanding, few could quarrel with Abelmann's interpretation and her graceful presentation. Her command of ethnography is a strength she brings to the book. Her dialogues, which reflect a new generation of Western "interpreters" sensitive to Koreans' misgivings toward these "others," open up discussions with various Koreans across class and regional boundaries. Such dialogues not only result in insightful interpretations of Korean culture and movement logics but also contribute to difficult and endless debates in Korea while resonating with the theoretical and conceptual concerns of those of us working in America.

The Abolition of Feudalism: Peasants, Lords, and Legislators in the French Revolution. By John Markoff. University Park: Pennsylvania State University Press, 1997. Pp. xviii+689. \$85.00 (cloth); \$25.00 (paper).

David Zaret

Indiana University, Bloomington

This book is a distinguished product of its author's long experience in subjecting that most contentious historical event—the French Revolution—to quantitative analysis. Markoff's goal is to explain the collapse of the seigneurial regime in terms of the ways in which peasants and legislators "confronted one another, posed problems for one another, and implicitly negotiated with one another" (p. 5). The quantitative analysis uses two sets of data. One is the *cahiers de doléances*, petitions that recorded grievances and demands as part of the process for convocation of the Estates-General of 1789. The content analysis of these grievances and demands uses the general *cahiers* endorsed by assemblies of the nobility and Third Estate and a national sample ($n = 748$) of *cahiers* from rural parishes (roughly 2% of the parish *cahiers*). The second data set has a longitudinal dimension that captures forms and rhythms of rural collective action up to 1793 (when the Convention passed legislation that ended seigneurial rights). From secondary historical sources, Markoff constructs a database of 4,700 "events" (20 or more persons who seize or damage resources of another party or defend themselves against the same).

Markoff's account of the peasant-bourgeois alliance against French feudalism opposes two dominant approaches to the study of the French Revolution. He rejects revisionist historical accounts that minimize the importance of long-term structural change and see elite manipulation as a principal factor behind popular participation in revolutionary politics. Yet his account is equally opposed to structural explanations that gloss over the importance of "process" (see below) in revolutionary politics. Throughout the book Markoff assesses and qualifies prior claims by social historians and sociologists on the political goals and tactics of different social groups. For example, evidence from the *cahiers* contradicts Alfred Cobban's claim that indemnification proposals (compensation for eliminated seigneurial rights) were a maneuver by nobles intent on preserving the seigneurial regime. Twenty-seven percent of parish assemblies supported compensation for at least one seigneurial right. That support varied widely for different seigneurial rights, Markoff argues, is evidence of a rational political calculus that led villagers to frame a viable political strategy. In analyzing rural political events, Markoff finds support for much but not all of Charles Tilly's portrait of traditional patterns of popular collective action: it was, as Tilly argues, predominantly "local" but it was not "patronized," that is, "neither made on behalf of nor rarely with the support of some powerful patron external to the community" (p. 263).

These specific comments illustrate three general themes. First, Markoff attributes more sophistication and rationality to peasant political attitudes than do many contemporary historians. Rural *cahiers* reveal "that French villagers show a thoughtful and nuanced capacity to differentiate among their burdens" (p. 72). Second, Markoff explores similarities and differences between grievances in the *cahiers* and subsequent legislative developments and forms of rural collective action. Variations in complaints in *cahiers* from different estates map onto subsequent political divisions in the National Assembly. Though rural *cahiers* are not precisely predictive of political developments, they shed light on subsequent collective action in the countryside. To explain the relatively small number of antitax revolts, Markoff cites hostility in rural *cahiers* toward seigneurial and ecclesiastical exactions. Grievances against taxes "are considerably more numerous" than those against exactions of the lords and church, but reform was more likely to be proposed for taxes, "outright abolition" for the exactions (p. 234). Third, Markoff emphasizes the "processual character of the situation in the countryside" (p. 411). The Revolution in the countryside "should be considered as a process" (p. 268, and see pp. 8, 13, 567) in order to understand how an antiseigneurial agenda, which did not dominate complaints in rural *cahiers*, became the focal point of rural insurrection. Process refers, among other things, to demonstration effects of one event for later events and specifically to "a process of bargaining" (p. 514) between revolutionary peasants and revolutionary legislators. Applied to the analysis of temporal rhythms for different categories of political events, this leads Markoff to the conclusion that "rural insurrections were at once causes, symptoms, and consequences of the breakdown of the political and moral authority of the Old Regime" (p. 288). What distinguishes this analysis from structural explanations in the tradition of Marx and Tocqueville—these, respectively, emphasize class antagonisms spurred by growing markets and a legitimation crisis caused by political centralization—is that "neither grapples with the Revolution itself as a political process" (p. 567).

Markoff covers many other topics of interest for scholars interested in the French Revolution or, more generally, in historical studies of political movements (e.g., spatial variation in collective action, literacy, political communication, revolutionary discourse on feudalism). Nonspecialist readers might find this diversity distracting. Several pages explain why the traditional preeminence of Sunday for popular political contestation became less pronounced during the Revolution. But Jack Goldstone's demographic-structural account of the French Revolution—"the most impressive analysis of the Revolution along these lines" (p. 566n)—is dismissed in a footnote. I wish Markoff devoted more attention to criticism (by Roger Chartier, e.g.) of the probative value of *cahiers* as evidence of peasant attitudes. These quibbles aside, this book must be read by anyone interested in the political sociology of the French Revolution.

The Rhetoric of Moral Protest: Public Campaigns, Celebrity Endorsement, and Political Mobilization. By Christian Lahusen. New York: Walter de Gruyter, 1996. Pp. xvi+425.

Mayer N. Zald
University of Michigan

The study of social movements has existed in a kind of time warp. Possibly due to the very success of the resource mobilization/collective action/political process paradigm, students of social movements have been slow to take the cultural turn that has so usefully swept some other parts of sociology and other social science disciplines. (Although the tide has begun to turn; see Hank Johnston and Bert Klandermans, eds., *Social Movements and Culture* [University of Minnesota Press, 1995] and Doug McAdam, John D. McCarthy, and Mayer N. Zald, eds., *Comparative Perspectives on Social Movements: Political Opportunities, Mobilizing Structures, and Cultural Framings* [Cambridge University Press, 1996]). Even more so, the insights and modes of analysis that have developed out of French deconstruction and postmodernism have not been applied to social movement phenomena. Comes Christian Lahusen, who wishes not to bury the currently dominant paradigm but to transform it and extend it by incorporating the important insights and analytic tools provided by semiotic analysis of signs and signifiers, of the meaning of rituals, and even, when appropriate, of the structure of musical composition.

Lahusen's focus is on four extended campaigns in which social movement organizations (SMOs) enlisted prominent popular musicians in international campaigns to raise money, attract media attention, attract new supporters and attach them to the movement, and contribute to political transformation. His four cases are (1) the Sun City album and the Artists United against Apartheid; (2) the Breakthrough and Rainbow Warrior albums and Greenpeace; (3) the Nelson Mandela birthday celebration and the British Anti-Apartheid Movement; and (4) the Human Rights Now world tour organized by Amnesty International.

The book is organized into five parts. Part 1, "Towards a Theory of Political Mobilization," introduces the reader to social movement theory, discusses the literature on and issues of extended campaigns, and argues that these campaigns have to be placed in their societal and institutional context. Part 2, "Investing in Popular Music: The Opportunities for Campaigning," describes in detail the institutional arrangements of the production of popular culture, especially popular music. The core issue is how do SMOs tie into institutional arrangements of record companies, copyrights, and so on to engage celebrity musicians in extended musical/political campaigns. To this point Lahusen's analysis is interesting, and the reader has learned much interesting detail about the music industry and the construction of campaigns, but new analytic ground has not been broken.

New ground is broken in part 3, "Designing and Composing Protest

Simulacra: The Campaign Event and Artifacts,” and part 4, “Understanding and Explaining Mobilization: Campaign Strategies and Organized Collective Action.” Popular music campaigns cannot use just any musician, nor any piece of music, nor any set of song words, nor any set of visual logos. They have to be fit to the specific messages that the SMO wants to convey. How are we to analyze these simulacra? Lahusen has to probe deeply into theories and methods of interpretation. Each form of simulacra has its own interpretative methods. For instance, the tools for analyzing the structure of music are different than the tools for analyzing the graphic designs of T-shirts, or posters. Moreover, campaigns are suffused with storytelling, not only by the media, but by SMO spokespersons and others. Thus, narrative analysis is needed to reveal the story lines and structures that are being emplotted. Obviously, each of these modes of analysis are subject to local epistemologies and must be judged within their own canons of evidence and interpretative rigor.

Moreover, Lahusen argues that these campaigns are in the business of producing reality, or a new hyperreality. Drawing on the ideas of Jean Baudrillard and Victor Turner, Lahusen shows how these campaigns attempt to create a political reality for the media, the larger public, and for participants in the mass rituals of campaign events, many of whom pay attention to the campaign not out of political interest, but out of sheer consumatory interest. They are fans of Peter Gabriel or Bruce Springsteen. Torture or environmental degradation may be the last thing on their minds when they buy a ticket to a concert or purchase a T-shirt.

Part 5, “The Globalization of Collective Action: International Campaigns in Context,” correctly notes that in two senses these campaigns usually involve an internationalized context—the SMO is involved in mobilizing support on issues that transcend single nations, and they do so in several nations. Thus, several issues in thinking about the targets of movements or the dilemmas of SMO operation in different national venues come to the fore.

The book is deep, serious, innovative, and provocative. On many topics, I found Lahusen to be penetrating. For instance, his discussion of micromobilization, nesting it in the campaign’s attempt to transform public representation and discourse, is rightly critical of rationalistic and individualistic notions of mobilization. Similarly, his analysis of signification processes and the ritual/performative aspects of campaigns is extremely illuminating. On the other hand, in its very completeness, the book is often tedious—in the presentation of so much theoretical and methodological material and so much analysis, the campaigns—their excitement, energy, and intensity—get lost.

Finally, this is a book that really could have used an index. I am sorry that the author or the publisher did not provide one.

Crisis and Transformation: The Kibbutz at Century's End. By Eliezer Ben-Rafael. Albany: State University of New York Press, 1997. Pp. xii+282. \$21.95.

Benjamin D. Zablocki
Rutgers University

This book is the result of a large-scale research project undertaken by Israeli scholars under the sponsorship of the Yad Tabenkin Research Center, the research arm of the Unified Kibbutz Movement. The data collection for this project, which was called "The Kibbutz at the Turn of the Century," was carried out between 1991 and 1994. Eliezer Ben-Rafael does an excellent job of making sense out of this huge amount of data as well as an even larger amount of statistical data that the Kibbutz movement has kept on itself for the years prior to 1991.

If Ben-Rafael had merely summarized the results of this massive data collection effort, his book would be, for that reason alone, a definitive and indispensable reference for all scholars interested in kibbutzim, intentional communities, or Israeli society as a whole. But the author has gone much further in contextualizing the transformation of the kibbutz in terms of the economic crisis of 1985 and the changing nature of the Israeli economy, the world economy, and the place of the kibbutz in both.

The book is divided into eight chapters that carry out a well-defined research agenda guided but never constrained by a wide range of theoretical ideas. In the first two chapters, the crisis facing the kibbutz is carefully defined. The next three chapters explore the dimensions of this crisis: the tension between community and individualism, the tension between economic rationality and socialist idealism, and the tension between the historical tradition of kibbutz elitism and the need to become better integrated within the very different Israeli society of the late 1990s. The sixth and seventh chapters go on to examine the agencies supporting and resisting change within the kibbutzim and the actual process of change as it has occurred among the kibbutzim since 1985. The final chapter relates all of this to the history of non-Israeli communal movements and speculates on what of enduring value may be salvaged from what the author convincingly shows is an inevitable trend toward individualism and increased economic rationality.

Although the accomplishments of this book are stunning, there are also a few disturbing omissions that detract somewhat from the overall achievement. The most striking omissions are in the area of ethnicity. Although the rapidly growing segment of the kibbutz labor force made up of outside wage laborers is discussed extensively in the book, the reader is left with the impression that all of these outside laborers are Jews who could have chosen to apply for membership in the kibbutz if they had so wished. Nowhere is it even hinted that some of these wage laborers are Arabs who are barred for ideological reasons from kibbutz membership.

Moreover, even in discussing the demographics of Jewish nonmember wage laborers, country of origin is not used as an independent variable. This is surprising because elsewhere in the book (chap. 6) the author shows important differences between kibbutz members of Sephardic as opposed to Ashkenazic origin in support of transformative changes in the direction of individualism and entrepreneurial capitalism.

The grudging and substandard care of the disabled elderly in some kibbutzim is another issue that is not even briefly mentioned in this book. This is surprising because one of the themes of the book is the increasing degree to which kibbutz members insist on developing professional identities outside the kibbutz rather than, as in the old days, giving their all to the community and trusting to be taken care of in times of need. Ben-Rafael ascribes the increasing reluctance to rely on the kibbutz for cradle to grave security to the economic crisis and the danger that some kibbutzim will be forced to disband in bankruptcy. While these are undoubtedly the most important factors, the attitudes of kibbutzniks themselves toward their elderly, particularly those whose children have moved away, should also have been discussed.

I do not want to give the impression that Ben-Rafael has glossed over the kibbutz movement's problems. On the contrary, this is a critical hard-hitting discussion of the kibbutz in crisis. This book is not at all a white-wash but is rather an honest critical analysis with a few blind spots. It would be helpful to complement this excellent work with another analysis of the same data by a team of scholars from outside Israel.

One final point is that, unfortunately, this important book is marred stylistically by poor copyediting. Although Ben-Rafael's writing is powerful and superbly organized, it is clear that English is not his customary language. SUNY Press should be ashamed of the minimal editorial attention they gave this book.

Historic Preservation: Collective Memory and Historical Identity. By Diane Barthel. New Brunswick, N.J.: Rutgers University Press, 1996. Pp. ix+182. \$48.00 (cloth); \$16.95 (paper).

Patrick Hutton
University of Vermont

Diane Barthel traces the rise of the historic preservation movement since the 19th century. She acquaints us with its growth and transformation from the beginning of the industrial age into our own postindustrial one, especially in England and the United States but with references to its recent globalization.

Barthel presents the project of preservation as the counterpoint to that of modernism, by which she means the making of urban industrial culture. The modernist vision opened upon the prospect of building a better future through science and technology. At the same time, the agrarian

past evoked a nostalgia that had to be accommodated. Ironically the makers of the new society, the industrialists themselves, played the pioneering role in funding historic preservation but always in ways that served their own interests. Barthel works with a number of topics that effectively cover the preservation field: the idealization of traditional society, the constructed image of industrialization, the commemoration of war, the preservation of religious shrines, and the globalization and especially the commercialization of the preservation movement.

In the process, Barthel offers a comprehensive analysis of what might be characterized as the politics of preservation, observed at the crossroads between memory and history. For her, that politics turns on the contest between a high idealism that would conserve the past in static form and a crass consumerism that would use it for present commercial purposes. Each presents perils by introducing elements that distort the reality of the past—idealization leads toward utopian thinking; commercialization leads toward sentimentalization. Pushed to an extreme, either of these purposes would thwart the quest of preservationists to represent the past with some measure of authenticity. The temptations to do so, however, are powerful, and studying their appeal is for Barthel what makes the history of the preservation movement interesting.

With respect to the shortcomings of idealization, Barthel launches her study with an account of the naïveté of the preservation movement in its early stages. In seeking to hold fast to the past they wanted to memorialize, preservationists turned it into a utopia exempt from historical change. Utopian settings are clean, harmonious, and serene, whereas the stuff of history is dirty, chaotic, and full of passion. In seeking to fix the past in timeless settings, moreover, they proposed a closure that was impossible to maintain. The past is a moving target, and in time the meanings sites of memory are supposed to evoke inevitably become dated. The tendency among preservationists sensitive to these trends, therefore, has been to recontextualize the places of historical memory in ways that make manifest tensions previously denied—an acknowledgment of groups excluded, conflicts that raged, issues unresolved. The quest to memorialize a complex past with some authenticity has in our times presented the challenge of designing sites that are open to multiple interpretations.

The problem at the other extreme, that of commercialization, is closely attuned to our postmodern times. Ours is a society caught up in the affluence of a consumer culture. Places of historical memory have become magnets for more and more tourists at leisure. Historical tourism has become big business, and those who design commemorative sites must consider the tastes of a mass culture. Hence curators of historic monuments are caught between the desire to educate and the need to entertain. If the problem of the early days of the preservation movement was its utopian idealism, that of the later projects has been their consumer-oriented banalization. The learning capacity of the average tourist is not that high, and the temptation toward entertainment is compelling. This Barthel refers to as "Disneyfication"—transforming heritage into kitsch that tugs

at human emotions in sentimental ways. Today it proceeds on a global scale.

Barthel concludes that the preservationists' search for authenticity is always an elusive quest, for preservation is an object lesson in the politics of culture. No single goal may be served. Caught between art and kitsch, the real and the ideal, high purpose and low commerce, education and entertainment, curators in our times have come to strive for balance. The past that is represented in places of historical memory always signifies a compromise. Historic preservationists may strive to present an authentic past, but their resources and their problems are different from those of historians, who theoretically have only other scholars to please.

Barthel is to be commended for having written a very useful book. She has done an outstanding job in contextualizing the issues preservationists face, yet she is never judgmental. In the process, she provides a perspective on the memory/history relationship that helps us to understand why it has recently inspired so much scholarly interest. For a better understanding of both the theory and practice of preservation, her book is a good place to begin.

The Spectacle of History: Speech, Text, and Memory at the Iran-Contra Hearings. By Michael Lynch and David Bogen. Durham, N.C.: Duke University Press, 1996. Pp. xi+348. \$49.95 (cloth); \$18.95 (paper).

Jane H. Hill
University of Arizona

As a linguistic anthropologist, I am committed to the idea that production and reproduction of culture and its product, "history," occur in the fine details of "discourse"—the moments of talk and written texts and encounters with them—and in the ways in which local discursive moments are articulated with larger systems by ideological and cultural filters on which discourses are evaluated as competent, meaningful, and effective. *The Spectacle of History* is an exemplary working out of such a commitment. First, it concerns an event that raised fundamental questions about the nature of civic life and state power: the interrogation of Oliver North in the 1987 U.S. Senate hearings on the Iran-Contra affair, an exchange that split its public into utterly opposed camps, one condemning North as a sleazy self-promoting liar, the other revering him as an authentic American hero. The authors undertake a detailed conversational analysis of several crucial exchanges in the hearings and the way that talk there articulated with the documentary record. But their analysis goes beyond ethnomethodological concerns (they include a detailed appendix on methodology) to raise major issues in social theory.

Bogen and Lynch argue that Oliver North and his supporters are "applied" postmodernists (or "applied deconstructionists"): "North . . . remorselessly asserts the right to advance his local knowledge of events"

(p. 72), disrupting the "master narrative," oriented to "reality," asserted by the Senate interrogators by showing how "their discourse was always already political" (p. 256). They argue that the idea of "master narrative" can be usefully reformulated to move away from the idea of a "totalizing mythology" (p. 73) advanced in much of the theory of postmodernity, to a "pragmatically respecified conception . . . that emphasizes the need for and use of a practical historical ground from which cogent narratives of past events can be assembled and ratified in situ" (p. 73). They exemplify this idea of master narrative as "practical" using the disputes over chronologies conducted during the hearings, where the opposing parties engaged conflicting documents in "a public contest over the writing of history, word by word, line by line" (p. 88).

As part of a very interesting inquiry into public "interrogation" as an institutionalized discursive form, Lynch and Bogen challenge Foucault's idea that such a "ceremonial of truth" is an efficient mode of production of rationality. They show that North was able to subvert this ceremonial, using the bureaucratic techniques of the "paper trail" and the minute inquisition of documents with exemplary skill to disrupt and mislead, while simultaneously performing as a "tragic hero" whose patriotic dreams were crushed by the dead hand of a bureaucratic Senate insisting on the letter of its constitutional role in foreign policy. In this discussion the authors open an important theme, the simultaneous importance in contemporary life of the "rational" techniques of record keeping and surveillance and the idea that public political talk should be oriented toward some testable reality, along with "more ancient forms of rhetorical and performative practice" (p. 95), where camera angle, a chestful of medals, and a full head of hair (the unruly strands of greying hair combed across the balding dome of the chief interrogator for the Democratic side of the Senate committee, Arthur Liman, contrasted unfavorably with North's made-for-TV haircut) are deployed with such success that North became seen as a serious Republican politician.

While talk is the usual site of ethnomethodological analysis (and one where Bogen and Lynch undertake innovative analyses of such phenomena as rhythm and pace and the moral dimension of "first-person" stories), Lynch and Bogen also pay close attention to the documents in the case, since these are the prototypical sources for history as traditionally construed. They point out how the Iran-Contra case confounds such a construal: Many documents before the committee were produced with one eye precisely on how they might "play" in such a hearing and the ways in which address lists on e-mail notes and paper memorandums would permit "plausible deniability" for principal actors in the affair. Bogen and Lynch develop many interesting analyses of the ways that the relationship between talk and text were manipulated in the hearings and how these reveal ideologies of discourse (e.g., in their discussion of the way that Liman tried to pry North away from his exasperating technique of spending long boring minutes with his lawyer consulting a briefing book in order to get him to talk "as a free subject" [p. 233]).

A brief review does not permit attention to the many interesting ideas in this densely detailed analysis, but I must say I took special pleasure in the proposal that *sease* become a technical term contrasted to *trust* (p. 246). I plan to recommend this book to colleagues and students as profound, useful, and wickedly entertaining.

Between Politics and Reason: The Drug Legalization Debate. By Erich Goode. New York: St. Martin's Press, 1997. Pp. x+181. \$35.00.

James A. Inciardi
University of Delaware

The drug legalization debate is now almost a decade old, having begun at the close of the 1980s when Baltimore mayor Kurt Schmoke, at some political risk, called for public discussions and consideration of radical alternatives to existing state and federal drug control policies. Mayor Schmoke and numerous other observers had long since decided that the more than seven decades of federal prohibition since the passage of the Harrison Narcotics Act in 1914 were not only a costly failure but a totally doomed effort as well. They argued that drug laws and drug enforcement had served mainly to create enormous profits for drug dealers and drug traffickers, overcrowded jails, police and other government corruption, a distorted foreign policy, predatory street crime carried on by users in search of the funds necessary to purchase black market drugs, and urban areas harassed by street-level drug dealers and drug gangs.

In the years since, there have been congressional hearings and government commissions focusing on the legalization issue, as well as hundreds of scholarly articles and books and perhaps thousands of media stories dissecting the opinions of everyone who has been publicly vocal in the debate. Given all of this discourse, one might ask if it was necessary for Erich Goode, a professor of sociology at the State University of New York at Stony Brook, to add *Between Politics and Reason* to the already lengthy list of drug policy material. As an author who has written much on the legalization debate, my answer to this question is yes. Goode's book is a welcome addition to the literature.

Much of what currently appears in the drug legalization literature is typically overopinionated, factually incorrect, poorly researched, or so one-sided that the reader is left with either a jaundiced or an impartial overview of the issues. Goode, who has been teaching courses on the legalization issue for more than two decades, has an excellent grasp of the drug literature, and, as a result, his book goes a long way to remedy this situation.

The book raises such questions as: Is it time to end the "War on Drugs"? Are we facing a "new crisis of legitimacy" in the criminal justice system when there are more inmates incarcerated today for drug viola-

tions than for violent offenses? Do these new and troubling developments cry out for drug legalization, or is there an easier, softer way?

Between Politics and Reason investigates these questions and examines the whole controversy surrounding the complex issue of legalization. Through an overview of drug use and abuse in America, a description of the prohibition model, and comprehensive analysis of today's key arguments, Goode's work offers a solid understanding of drug legalization, the harm reduction approach, and the war on drugs. The book also considers how tobacco and alcohol—possibly the most harmful of all drugs—could be controlled.

For those of us who have participated heavily in the drug legalization debate, there is little in *Between Politics and Reason* that we have not heard already. Moreover, the great majority of the debaters (on both sides of the issue) are so committed to their positions that even this well-written book will likely not change their minds. However, there are scores of others who need a balanced treatment of the issues, and this book may just be the answer.

Hybrids of Modernity: Anthropology, the Nation State and the Universal Exhibition. By Penelope Harvey. New York: Routledge, 1996. Pp. ix+208. \$65.00 (cloth); \$22.95 (paper).

Victoria D. Alexander
University of Surrey

In the Expo 92, Switzerland mounted a postmodern exhibit in its pavilion. They invited 12 artists, not all Swiss citizens or residents, to create installations. Rather than celebrate the concept of the nation-state, most artists chose to question it: A musical piece blended 96 voices each speaking a distinct Swiss dialect. Another work, declaring "La Suisse n'existe pas," defined the nation with the Descartesian pun, "Je pense donc je suisse." The nation-state, a prototypically "modern" institution, has been bolstered as a concept through world's fairs since 1851. How does "reflexive" questioning, like the Swiss pavilion, affect the nation-state? What happens to the notion of the nation-state in an era when "nation," "state," and "country" (i.e., culture, government, and territory) no longer overlap and institutions like the Expo refuse to reify it? These are some of the questions Penelope Harvey poses in her book, *Hybrids of Modernity*.

Many nation-states took a conventional, modern approach to their displays, concentrating on technology, progress, stability, continuity, and coherence. Chile, for instance, brought a large piece of Antarctic glacier to demonstrate that it had mastered the technological prowess such a feat entails (and thereby angering environmentalists who were disturbed by the appropriation of icebergs). The U.S. pavilion focused on the Bill of Rights by showing, as Harvey ironically phrases it, "a film presenting

the inalienable human rights of liberty, life, and the pursuit of pleasure" (p. 66).

Some countries tried to combat stereotypes. Japan chose to focus on tradition, rather than technology, by building its pavilion of wood and eschewing displays of nifty new consumer gadgets. But this approach disappointed visitors who expected Japan to show the latest electronic objects. The lack was especially noticeable in contrast to the Fujitsu corporate pavilion, located next door, whose theme was progress through technology. Stereotyped misunderstandings abounded, as when a couple took offense at the "American Spirit House." They thought that the message intended by the bathroom showcase was that arrogant Americans did not think Spaniards had indoor plumbing.

Most visitors did not have time to absorb all the proffered messages. Running from pavilion to pavilion with an Expo 92 "passport," which was stamped by each country at the exit of each pavilion, was not conducive to reflection, and, indeed, probably heightened stereotypes. As Harvey points out, "For many people, the accumulation of these signs [the passport stamps, also VIP pins, various freebies, and purchased souvenirs] was more important than the content of the exhibit itself" (p. 156). Audiences came to the Expo 92 for entertainment and spectacle, not for education. The most popular exhibits used film technology to provide amusement-park experiences, and the lines to get into these exhibits were enormous—often eight hours or more in the sweltering Seville sunshine: "There was intense competition to get to the front of the queues, people literally running across the site as soon as the gates opened in the mornings. Many others decided on an alternative strategy, to see as much of the rest of the exhibition as possible. Some . . . gave up and sat in the ornamental pools" (p. 155–56).

Harvey suggests that the presence of both modern and postmodern forms of representation in Expo 92 shows that the Expo, as an institution, has become a hybrid form. A mixture of modern and postmodern forms also characterizes nation-states. Visitors were viewed as modern citizens encouraged to identify with their country, and, simultaneously, as postmodern consumers who had a choice of what to believe, identify with, or find pleasurable. Further, the Expo treated nation-states and multinational corporations as equivalent, giving pavilions to each, further blurring the modernist idea that corporations are part of the technological force of nations.

Expo 92 provides a rich vein to mine, and when Harvey gives concrete illustrations or interpretations from the event, they are invariably intriguing. Still, I could have done without the plethora of sentences like: "The constructed forms of the Expo are the materiality of communicative practice, indeterminate signs dependent on interpretive outcomes, themselves situated within parameters of possibility which are entirely cultural or contingent even when self-evident to the parties concerned, even when tightly confined in apparently over-determined webs of signification" (p. 173).

Hybrids of Modernity, however, is not primarily about the exhibition. Rather, it is a rambling reflection on anthropology and how it is changing in the postmodern era. Harvey wishes to show that cultural studies and anthropology are compatible. But what she does instead is inadvertently demonstrate the pitfalls involved in taking a postmodern tack. How can a discipline like anthropology, whose traditional mission has been to interpret cultures, do so in an era that recognizes multiple interpretations and decries the interpretations of the powerful about the weak? Harvey's solution is "auto-anthropology," studying ourselves. This, of course, poses its own problems requiring anthropologists to closely examine themselves, their mission, and their discipline, preferably with a heavy dose of jargon. What can this book teach sociology? I am sorry to say, not much. Like postmodernism itself, Harvey's book raises more questions than it answers and it is just as hard to pin down.

Interaction and Grammar. Edited by Elinor Ochs, Emanuel A. Schegloff, and Sandra A. Thompson. Cambridge: Cambridge University Press, 1996. Pp. xii+468. \$29.95 (paper).

Allen D. Grimshaw
Indiana University

Interaction and Grammar is an innovative, informative, interesting, and, in my view, important book. My principal disappointments are that I have insufficient space to exemplify its full range of value for sociologist readers (or to discuss individual contributions), that the special parlance employed by contributors may discourage some potential readers, and that some apparent shifts in the direction of conceptualization and research in conversation analysis have not been pushed further—and made more explicit.

The book is innovative in that the collaboration it reports is among pioneers in articulating work in formal syntax, conversation analysis, and paralinguistics (e.g., nonexhaustively—voice quality, intonation, and tempo), kinesics (e.g., nonexhaustively—facial expression, gesture, and "body language"), and proxemics (interpersonal distancing and arrangements of "things" in space) in investigation of interactional accomplishment. It is informative in illustrating something of the range of autonomously syntactic (e.g., clausal organization and reorganization, employ of inflection and particles, and movement of nouns or verbs), prosodic and paralinguistic, pragmatic, and other devices available in different languages/cultures (English, Finnish, Japanese, Kaluli) for management of turns (including, e.g., shifts of rights to the floor, acceptance or rejection of interruptions, simultaneous speech, etc.), for making credible claims, or cautious characterizations of those not present, or accomplishing joint response to disruption. Fox, Hayashi, and Jaspersion claim that English clause beginnings are richer than Japanese with information

about "how clause is likely to continue" thus providing potential interrupters with valuable information. They also observe that repair (by speaker or other) extends possibilities for how an utterance can be completed *in any language* (p. 220). They are talking about *syntactic resources for interaction*. Some of these resources/devices are both familiar and fun in a Goffmanesque manner as when we are introduced to, and immediately resonate to, such notions as "trailoffs" and "rush-throughs" (Schegloff) and "outlastings" (Lerner).

The book is informative because, in addition to demonstrating the interinfluence of syntax and interaction, it provides windows to sociological understandings revealed when seldom exploited perspectives are applied to concerns at the core of sociology including, illustratively, issues of symbolic interaction, of the socialization of neophytes, and of social change. Beyond this it allows similarly (and seminally) novel views of such specialties as the sociology of science and that of occupations. I found particularly instructive, for example, Ochs, Gonzales, and Jacoby on how scientists construct indeterminate referential identities—sometimes in the process blurring the distinction between themselves and the physical world under their scrutiny—and how meaning is built through routine interpretive activity involving talk, gesture, and graphic representation. Other readers will perhaps find more engrossing Goodwin's rich integration of different channels of behavior (syntactic production, intonation, body movement, display of awareness of a world beyond the immediate and ongoing) in a more dynamic reconceptualization of Goffman's notion of participation framework as it operates in situations of disruption (of initially unknown magnitude) of routines—in an airport control tower.

The book is interesting because of sociologically relevant questions posed, answers given, and passing observations. Lerner, for example, continues here his thoughtful and suggestive work on how and why interlocutors complete utterances of speakers without being asked initially suggests as functions of such completion (1) agreement, (2) preemption of disagreement, (3) collaboration, and (4) heckling (p. 244). He later argues that an early opportunistic completion may be intended to initiate or sustain a special alignment with a speaker such as affiliation (pp. 263–64). Sorjonen suggests that repeats and the Finnish particles *niin* and *joo* variously function as (1) interrogatives, (2) exclamations, (3) requests for confirmation (pp. 279–80), (4) challenges, and (5) expressions of ritualized disbelief. There are obvious parallels to English. She also has some very nice observations on sweetening recommendations of others when the recommender suspects an unwanted invitation or request to self may be forthcoming. Schieffelin shows how the invention and introduction of an evidential construction to refer to *printed* religious material, translatable as "known from this source/not known before" has not only granted authority to written text when there is no basis in fact for so doing but has also been associated with the introduction of higher status for a new role of interpreter of Christianity in a society where prior stratification rested on quite different bases (and, not incidentally, also to a lowering of the

status of women in a previously more egalitarian society). I wonder what the missionaries will think of Schieffelin's characterization of what they are doing (p. 439).

The book is important, I believe, for its general demonstration of the reciprocal necessity and value of sociological and linguistic methods and perspectives—and for the specific instances noted above. I am increasingly hopeful about a rapprochement of disciplines sharing a common (but sometimes neglected) interest in language in use in social contexts. It is important, too, because it has foregrounded often neglected questions about incorporation of prosodic/paralinguistic phenomena in studies of social interaction (and emotion) and grammar (and social semantics). Finally, I can only hint at how important I think it is that Schegloff and other conversational analysts are gradually moving toward more explicit acceptance of situational/historical context as relevant for full understanding of what goes on in talk. Schegloff, Ochs, and Thompson and Schegloff in his own chapter appear to suggest that conversational analysts have always known this and acted on that knowledge. I have a strong sense that for Schegloff, minimally, context still means those of text and sequence more than of history, biography, and the larger world. I will read forthcoming studies influenced by this important volume to watch the broader view gain wider acceptance.

Those who know me and who read the book will know that I disagree with various of its claims, interpretations, views of history, and so on. They will also understand why I choose to use my limited space to emphasize the volume's many positive features.

French Intellectual Nobility: Institutional and Symbolic Transformations in the Post-Sartrean Era. By Niilo Kauppi. Albany: State University of New York Press, 1996. Pp. xii+204. \$21.95 (paper).

Charles Lemert
Wesleyan University

In the mid-1970s, I spent a year in Paris. Each morning, I took my youngest child, then just two years old, to a bilingual nursery school. Each morning, we waited for bus 69, which stopped at the corner of Rue Saint Jacques and Rue des Écoles. Each morning, bright and early, Michel Foucault walked past us on the way to the Collège de France. I never spoke to him, but I could have. Needless to say, others did. Such is life for the intellectual nobility in France. One is constantly at risk of a personal encounter with one's admirers or enemies.

Among many of the interesting observations in *French Intellectual Nobility*, a small but well-packed book, Niilo Kauppi reminds us that the French in particular do their intellectual work under conditions utterly different from those in other places, especially places like America, where it is entirely possible to say whatever one wants about another's ideas

and to speak under the illusion (and it is an illusion) that the potential enemy will never confront us face-to-face. In France, only the foolish believe they will never encounter an intellectual opponent at the bus stop, in a café, or at the cafeteria of the *Maison des Sciences de l'Homme*.

As a result, French intellectuals must always do their public work with an eye on their personal, even private, lives. Kauppi, for example, points out that the writings of his teacher, Pierre Bourdieu, are sometimes double voiced, as though "the narrator, the individual with a body, is separated from the narrator's sociological identity" (p. 94). Those who have not submitted themselves to long years of patient exposure to French writing and thinking are typically (and understandably) annoyed upon reading Bourdieu, Lacan, Foucault, or Derrida. Foucault is regularly criticized for failing to name and cite his references. Derrida, it need not be said, is famous for his play with words and names. Lacan's ultimately simple ideas are lodged in dreamlike abstractions. And it is surely true that Bourdieu's writing is very hard on English and American readers who are accustomed to more plainspoken discourse.

French Intellectual Nobility, though it has numerous and serious academic aims, ultimately succeeds mostly for its ability to introduce the reader to many interesting sociological facts accounting for the remarkable and globally important transformation of French culture and thinking since the days, just after the Second World War, when Jean Paul Sartre was the symbolic embodiment of what the rest of the world recognized as a French intellectual. Specifically, the book's focus is upon "the sociohistorical conditions for the production of theory in the French intellectual field in the post-Sartrean era, roughly since the end of the 1950s" (p. 4). Its author, Niilo Kauppi, is a research fellow at the Academy of Finland, which fact alone well suits him to the task. The Finnish people speak a language in the Finn-Urgic group that spreads from Norway and Finland in the west to Siberia in the east. They, thus, think in a tongue that bears little natural affinity to the languages of world domination and are accustomed to learning the discursive ways of others. Kauppi, for example, in addition to his obvious fluency in English, was educated in French as well. This allows for his interest in and grasp of his subject—and for his ability to interpret the French to English speakers.

The book is organized into three sections: (1) The French Intellectual Habitus and Literary Culture, (2) The Rise of Structural Constructivism and Semiology, and (3) Iconoclasm and Parody: The Literary Avant-Garde against Sartre. In each instance, the influence of Bourdieu is apparent—most especially in the longer second part, which is principally an interpretation of Bourdieu's distinctive sociology. Kauppi's argument is, roughly, that upon the decline of Sartrean existentialism there arose, as we know, a structuralist correction. As important as Lévi-Strauss and Foucault were to this correction and the boost it gave to the human sciences in France, the chief problem with French structuralism and post-structuralism was its lack of an explicit method for the analysis of social

institutions. The claim is made that Bourdieu's idea of the habitus (borrowed from Durkheim, Norbert Elias, and Erwin Panofsky) has provided just that "missing link" (p. 5). There is much to be said for this argument, even if its importance may well be lost on English and American readers more accustomed to the near opposite sociological deficiency: an institutional analysis that only recently has begun to consider cultural and symbolic matters.

Bourdieu himself is famously edgy about unauthorized interpretations of his ideas. Whether he approves of this one by a former student, I cannot say. I found it thorough and generally reliable, even though it sometimes states its case rather too formalistically (as when Bourdieu's use of the concept *capital* is presented as though it might be an economic, as distinct from metaphoric, usage). Yet, I have no doubt that readers who are not already familiar with Bourdieu and the French in general could learn quite a lot from this book, especially in its informative explanation of the place of social theorists like Lévi-Strauss and Bourdieu in the larger field of French culture. The third section, for example, offers an interesting, if somewhat tangential, discussion of the Tel Quel group.

This book is packed with curiosities. Kauppi reports (p. 128), for example, that between 1981 and 1989 Bourdieu's ranking (by the French) among France's most influential intellectuals rose from thirty-sixth to ninth. This was the period during which Roland Barthes, Lacan, and Foucault died. The implication is that, with their deaths, Bourdieu's ideas rose to provide the missing institutional link the earlier structuralist or semiological thinkers underplayed. This is an idea well worth entertaining. Experts on French intellectual history will not learn much new from *French Intellectual Nobility*, and they will surely quibble here and there with details. But those new, or relatively new, to the subject might well be entertained and informed.

Fanon's Dialectic of Experience. By Ato Sekyi-Otu. Cambridge, Mass.: Harvard University Press, 1996. Pp. x+276. \$39.95 (cloth); \$19.95 (paper).

David Scott
University of the West Indies

From about the middle 1960s through the early 1970s, Frantz Fanon was a symbol of the insurgent and uncompromising desire for a total revolution against colonialism. *The Wretched of the Earth* (published shortly before his death in 1961) is the master text of this moment. Around it there grew a complex debate about Fanon's understanding of colonial power and the process of decolonization, about the role of the peasantry and lumpen proletariat, about the place of violence, about his characterization of the national bourgeoisie, about his relation to European humanism, and so on. In short, to think through the question of the decolo-

nization of the Third World in these years one had *to pass through* Fanon's revolutionary thought.

As political independence became the accepted norm, however, and the hard years of the making of the "new nations" displaced the epic vision of the liberation struggle, Fanon's name faded somewhat from the frontline of Third World intellectual concern. In the middle to late 1980s, Fanon was rediscovered, not in the Third World, though, but in the metropolitan diaspora. If the 1980s are "lost years" in the Third World, they were years of a certain cultural intellectual awakening in the North Atlantic postcolonial diasporas. What is taking place here is a shift in what counted as the authoritative site of the *contestation* over the problematic of "the colonial," a shift at once geopolitical and philosophical. And the text through which Fanon is reread in this new space of diasporic displacement is not so much the manifesto of Third World revolution, *The Wretched of the Earth*, as the earlier meditation on identity, *Black Skin, White Masks* (originally published in 1952).

Today, there is a growing skepticism regarding the purchase of the critical strategies through which (and the cultural political project into which) this Fanon has been read. It is into this emerging space that Ato Sekyi-Otu's engaging book, *Fanon's Dialectic of Experience*, is inserted. Sekyi-Otu is one of a number of critics (see also Lewis Gordon, *Fanon and the Crisis of European Man* [Routledge, 1995], and Lewis Gordon, T. Denean Sharpley-Whiting, and Renee T. White, eds., *Fanon: A Critical Reader* [Routledge, 1996]) who have doubts about the focus of the postcolonial critics who have been responsible for a large share of the recent interest in Fanon. Sekyi-Otu is not unimpressed by this work, but he has two sorts of criticisms. In the first place, he departs, he says, "from their overwhelming concentration upon the significance of Fanon's work for understanding the psychodynamic of otherness and marginality." He wishes to reinstate a more *politically* engaged Fanon. In the second place, if much the better part of the work on Fanon written by postcolonial critics has operated on the North Atlantic diaspora black experience (one thinks here of the work of Homi Bhabha and Stuart Hall, e.g.), Sekyi-Otu is exercised by what he calls "the disasters of the post-independence experience in Africa." This is the geopolitical/intellectual space of Sekyi-Otu's intervention.

Now whether or not you agree with Sekyi-Otu's characterization of the position of the postcolonial critics (and I do not entirely agree), there is a dimension of his dissatisfaction that in my view demands a considered response—this is the problem of the *location* and the *register* of the target of the critique of postcoloniality. *Fanon's Dialectic of Experience* has the virtue of pointedly formulating this question. At the same time, however, it seems to me that an adequate account of this has to do more than merely chastise postcolonial critics for their neglect of politics and the predicament of the Third World as a privileged object of criticism. Such an account has also to provide a genealogical description that connects

the successive geohistorical *demands* on postcolonial criticism (including that to which Sekyi-Otu himself is responding) to the shifting conceptual institutional as well as geopolitical problem spaces in which the Third World as an object of knowledge is produced.

The most provocative argument in this book concerns its attempt to dislodge that reductionist reading of Fanon that treats his texts as conclusive propositions. Sekyi-Otu argues (convincingly to my mind) that although Fanon's interpreters disagree about much in their reading of his work, they "are virtually united in according utterances in his texts the unambiguous status of propositional statements and doctrines" (p. 3). The emblematic case, as he says, is Fanon's discussion of violence in *The Wretched of the Earth*. Those who agree and those who disagree with Fanon "have found in this discourse a doctrinal prescription" (p. 3). Against this, Sekyi-Otu proposes an alternative hermeneutic procedure, namely, that "we read Fanon's texts as though they formed one dramatic dialectical narrative" (p. 4). This means that the relationships between utterance and proposition, representation and truth, cannot be taken to be transparent. It means that an utterance or a representation encountered in a text is to be considered not as a discrete and conclusive event but as a strategic and self-revising act (p. 5).

Less convincing is the overall project of a reading that hopes to recover/recenter "the universal" in Fanon's work. In his view, the "signal feature" in Fanon's oeuvre is "the irrepressible presence" in his critical vision "of an openness to the universal." The restoration of the project of the universal is crucial to Sekyi-Otu because he is suspicious of the politics of difference in postindependence Africa. For him, among the "harrowing lessons" of this postcolonial experience "are the uses and abuses of the plea of particularity." "In the postcolonial world," he says, "'the night of the absolute' . . . became the nightmare of absolutism" (p. 20). It is easy to see what Sekyi-Otu's worry is. I for one am in sympathy with his political hopes. But I have my own doubts about this return of the universal. It seems a bit spurious, at the least, to conflate the ethics of difference embodied in the work of the critics of colonialist discourse with the authoritarian politics of ethnic genocide in postindependence Africa. Perhaps we need another way to go about negotiating the relationship between an ethics of difference and the demands of a politics of transformation.

But let us say that you are persuaded by the provocative story offered in this book. Let us say that it would be hard to deny that Sekyi-Otu has gotten *his* Fanon right, that he has illustrated beyond doubt the poverty of the poststructuralist reading, that he has convincingly rescued the critical salience of the humanist openness to the universal, and that he has admirably reestablished to Fanon's texts the integrity of a revolutionary interpretation: what then? What exactly is the yield of this strategy of criticism as an historical criticism of the postcolonial present? Is it enough to get Fanon right? Or is there another, larger, question to be raised, a question about the *horizon* in relation to which a Fanonian politics

is articulated? In other words, how do we stand in relation to the anti-colonial project of Fanonian *futures*?

Modernity and the State: East, West. By Claus Offe. Cambridge, Mass.: MIT Press, 1996. Pp. xii+270. \$36.00 (cloth); \$18.00 (paper).

Fred Block

University of California, Davis

In the 1980s, Claus Offe published two influential books of essays, *Contradictions of the Welfare State* (MIT Press, 1984) and *Disorganized Capitalism* (MIT Press, 1985). The present volume follows closely the format and the subject matter of those earlier books. *Modernity and the State* brings together translations of 12 of Offe's essays that were originally published between 1986 and 1994. The core subject matter is the deepening crisis of European welfare states.

Offe defends the accomplishments of the European welfare states, and he has little patience for the simplified arguments of neoliberals who insist that free markets will solve all problems. However, he also believes that the programs and demands of the social democratic left are largely exhausted and that purely defensive struggles to hold on to earlier gains are politically futile. Offe is particularly cogent in his explanations for the rising political suspicion toward state action. His search is for a third path—innovative strategies that could contribute to greater equality and greater democracy—without the negative baggage of fiscal crises and overreliance on an overburdened state.

In two essays at the beginning of the volume, Offe seeks to provide a deeper diagnosis of the current crisis. While later essays invoke the usual suspects in the crisis of the welfare state—greater global integration, the contraction of industrial employment, and diminishing social homogeneity—the emphasis here is placed on rigidities that are inherent in modernity itself. The argument is that precisely those social arrangements that are designed to maximize individual options and freedom of choice have the consequence at the aggregate level of making society more rigid and less able to adapt its institutions. The analogy is with automobile transportation: the elaborate network of roads and highways gives individuals unprecedented options, but the cost to society as a whole is an enormously inflexible investment in highway infrastructure that, in turn, constrains choice in other social arenas.

One aspect of this institutional rigidity is an ongoing deficiency in "steering capacity" in the political realm. Governments lack the resources—in budgets and political legitimacy—to cope with such large-scale coordination problems as "military-strategic questions of peacekeeping, the management of problems of poverty in the Third World, and the maintenance of minimal ecological balances" (p. 15). But even less daunting tasks, such as revising a pension system, become increasingly hopeless

given the fragmentation of interests within the political sphere and the heightened vulnerability of particular groups.

In short, Offe understands the economic and political crises of the welfare state as flowing from the "iron cage" of modernity; our very success in expanding individual options has made our societies progressively less governable. Since he does not believe that this problem can be solved either by greater reliance on the market or by increasing the capacities of the state, his preferred choice lies in a strategy of "unburdening" society's steering mechanisms. The way to do this is through decreasing social interdependence so that the need for overall coordination and steering will actually be reduced.

In the early essays, this unburdening strategy is addressed at a high level of abstraction, but Offe's meaning becomes clearer in some later essays that deal more concretely with questions of social policy. One chapter, written with Ulrich Muckenerger and Ilona Ostner, makes the argument for a system of basic income that would be guaranteed by the state and that would replace many existing transfer payments. The level of support would be high enough that recipients were not impoverished, but they would still have a financial incentive to take paid work. Neoliberals would be quick to dismiss such a proposal as a huge government spending program that would further burden the state, but Offe sees this idea as fitting with the logic of unburdening.

The system of basic income would replace the current elaborate system of unemployment insurance, which in Germany has been based on the assumption that the economy will usually be close to full employment. As a consequence, the unemployment insurance system has been subject to enormous strains during years of chronically high unemployment. The shift to a transfer regime based on more realistic labor market assumptions should make it easier to steer the society. Second, by effectively separating work and income, the basic income system makes individuals better able to cope with disruptions in their earnings from work and diminishes their vulnerability to small shifts in government or business policy. The result would be to loosen the coupling between the political system and the economy. Third, the availability of income would free an increasing number of people to work for cooperatives and neighborhood and self-help organizations that are able to provide services without direct reliance on government. The result could be both greater individual choice and a reduction in society's deficit of steering capacity.

While some of these essays are difficult and some of the translations could be more felicitous, the rewards to the diligent reader are considerable. Offe is one of our most powerful and original social thinkers, and his work directly illuminates the political and economic dilemmas faced by the developed societies of Europe and North America.

Realignments in the Welfare State: Health Policy in the United States, Britain, and Canada. By Mary Ruggie. New York: Columbia University Press, 1996. Pp. xiv+298. \$17.50.

Terry Boychuk
Macalester College

Realignments in the Welfare State seeks to identify three successive stages in health policy formation in the United States, Britain, and Canada. Social policy in these countries presumably exemplifies a liberal democratic, as opposed to social democratic or corporatist, approach to welfare state development. Ruggie asserts that these nations first saw the emergence of what she terms "segmented" policy regimes, they were followed by "interventionist" ones in the 1970s and 1980s and now may be witnessing the unfolding of a third, so-called "integrative" policy regime. Segmented policy regimes in this analysis correspond to the classical notion of liberal democratic politics, where public policy observes and upholds the separation of state and civil society, granting to each individual independent and differentiated rights, responsibilities, and prerogatives. In health policy, segmentation characterized the rapid growth of the public funding of health services in the immediate postwar era. Governments preserved to the greatest extent the traditional autonomy of health care professions and institutions while the latter generally acceded to the state's obligation to enlarge the scope of the public fisc to expand access to health care. Interventionist policy regimes grew out of the fiscal stringencies of the past few decades. Budget-conscious governments in these three countries unilaterally intervened in the health care industry, seeking to reconstruct health care provision in various ways so as to moderate spiraling costs. These efforts provoked resistance for violating the customary independence of health care providers, and they ultimately failed. Overcoming the stalemate accompanying interventionist policies may provide the impetus for the creation of an integrative policy regime, according to Ruggie. Integrative policymaking, distinct from segmentation, elevates collaboration and mutual accommodation among public and private actors over observing a rigid division of labor between them. Governments, distinct from interventionism, relax their use of regulatory power over health care providers in favor of voluntaristic discourse, persuasion, and consensus building around the ultimate objectives of policy.

The thesis of *Realignments in the Welfare State* is an interesting one; however, the conceptual and empirical framework of the study does not lend conclusive support to the author's principal argument. Since Ruggie claims that segmented, interventionist, and integrative policy regimes are exclusive to liberal democracies, it is worthwhile to revisit Gösta Esping-Andersen's typology of welfare states upon which this book is premised. Esping-Andersen's classification of liberal democratic, social democratic, and corporatist welfare states links specific, complex historical political processes with particular policy outcomes. The distinctive programmatic

features of social policies are the degrees of decommodification (universal vs. targeted entitlements) and of stratification (equal vs. differentiated benefits) explicit in their design. The policy outcomes referenced in this schema were primarily income maintenance programs. Health policies were not considered in the elaboration of the typology noted above, and, most commonly, students of welfare state development have consciously situated health policy beyond the scope of more general theories of social policy making. There is no theoretical warrant to draw inferences about health policy from this classification of welfare states, as the author does here, while there are obvious and compelling empirical reasons to treat health policy as following separate trajectories from those identified in generalizing accounts of social policy formation.

Canada and the United Kingdom have policies that strive to achieve universal and equal access to health care. Any interpretation of Esping-Andersen's criteria for classifying social policies could only identify the thrust of health policy in these countries as social democratic, not liberal democratic, as is apparent in the United States. This casts serious doubt on Ruggie's central claim about the essential comparability of American, Canadian, and British health policy. Even if readers were to suspend their disbelief about the commonality of health policy in these countries in terms of the *outcomes* of political contention and to focus on political *processes* as a unifying phenomena, as is suggested here, then the author would do well to establish comparisons with social democratic and corporatist states demonstrating the uniqueness of liberal democratic processes governing health policy formation. The study does not make this necessary argument. The distinction can no longer be assumed. There is, however, ample reason to suspect that social democratic, corporatist, and liberal democratic states have much in common in their relations with well-organized medical professions bent on maximizing their economic and clinical autonomy.

Setting aside this criticism for the moment, Ruggie may have still captured some important unifying trend in the political process surrounding health policy in these three countries notwithstanding questionable claims about their unique relevance to liberal democracies. The author's overview of historical and contemporary health policy in the United States, Britain, and Canada does not, however, give the reader sufficient evidence to consider the proposition as proven. Ruggie acknowledges that segmental, interventionist, and integrative tendencies are not mutually exclusive. They appear in all policy deliberations, albeit in varying degrees. As such, the author can only demonstrate which tendency prevails over the others in policy making with the aid of unambiguous operational definitions of these concepts. The organization of the empirical chapters of the book gains no assistance from the elaboration of transparent, rigorous tests for classifying policy episodes as primarily segmental, interventionist, or integrative. The author makes these judgment calls without reference to consistent, substantiating criteria. The thesis awaits a much more carefully articulated proof.

Big Business, Strong State: Collusion and Conflict in South Korean Development, 1960–1990. By Eun Mee Kim. Albany: State University of New York Press, 1997. Pp. xvii+280. \$19.95.

Changwon Lee
Korea Labor Institute

The active role of the state in economic development has recently been questioned by scholars as newly industrializing countries' goal to catch up is being realized. As a consequence, research interests have dramatically changed from state-centered approaches and the strong state model to the state-business alliance model specifically and the state-society relationship in general. States are no longer treated as the dominant independent variable in examining the recent interplay among institutional actors. However, it is still a lacuna how and in what way the developmental state in newly industrializing countries has been transformed or declined and how the state-dominated political economy in those countries has shifted into an alliance role with society.

In this context, Kim's book, which delves into the case of South Korea, provides an important angle for explaining the rise and decline of the developmental state. According to Kim, it is basically due to the inherent and contradictory nature of the developmental state that the state limits its intervention in the economy after the state completes its goal of economic development. Particularly in South Korea, where the state's developmental strategy fostered large businesses during the past three decades, the state paradoxically has recently encountered a situation in which the large businesses often override the state's decisions. As the state attains economic development, its nature changes from a comprehensive developmental state to a limited one. The main contribution of this book is positioning the nature and role of the developmental state in a dialectic and dynamic process of social change, overcoming the traditional state model of the strong state, which often simplifies.

This book is organized into two parts, in addition to the introduction and conclusion. The first part addresses the two key institutions, the state and the *chaebol*, that brought about rapid economic development in Korea. In treating these two institutional actors, Kim attempts to demonstrate that both parties have consistently been relatively autonomous from the other. In other words, from the beginning stage of economic development, the relationship between the state and business was one based on alliances and cooperation, contrary to the widespread notion that the state dominated the relationship.

The author argues that when the state in newly industrializing countries is examined, the two orientations of the state, developmental and authoritarian, should be analyzed separately. According to Kim, Korea's transformation from a comprehensive developmental state to a limited one, though it basically resulted from the inherent contradictions of the developmental state, was compounded by dramatic democratization in

the 1980s. However, in the following section, it appears that Kim provides little evidence for unraveling the compounded relationship between democratization and the decline of the developmental state in South Korea. Rather, some momentous changes in South Korea may be interpreted as caused mainly by the collapse of the authoritarian regime. Thus, it is arguable that the phenomenal transformation from a comprehensive to a limited developmental state was mainly driven by the demise of the authoritarian state.

The second key institution, the *chaebol*, is depicted as a family-owned and family-managed industrial organization influenced and affected by its environment (i.e., the historical and structural context) that at the same time takes advantage of and even changes its environment. As a result of institutional isomorphic change affected mainly by the role and policies of the state, the *chaebol* became the model of an industrial organization that other businesses tried to emulate. The causal linkage leading to the organizational isomorphism of the *chaebol* is, however, tenuous in the book. Big and diversified businesses are not peculiar outcomes influenced by the developmental state. The only aspect that distinguishes South Korea's case is that the *chaebols* were family owned and possessed a family-managed governance structure. Family ownership and management are, however, organizational attributes that are less vulnerable to the influences of the state. Rather, it is an institutional attribute of Korean society, which is based on familial ties. As a result, I could hardly find evidence in this book that the organizational attributes of *chaebols* were mainly affected by the state, while it confirms the historical truth that *chaebols* could grow under the influences of the state.

In part 2, Kim argues that the relationship between the state and the largest *chaebol* in the 1970s was more interdependent than the state centrists argued. The interdependent relationship changed into a more independent relationship by undergoing the decline of the developmental state and the rise of the *chaebol* during the 1980s. According to the author, "this is because capitalist economic development will inevitably result in the growth of the capitalists, who can then threaten the autonomy of the state" (p. 18). Although she emphasizes a dynamic model in addressing the relationship between the developmental state and the *chaebol* in her analysis, Kim subsumes this model under the static notion of the inherent contradictions of the developmental state.

In sum, although it is doubtful that the dynamic model of the state-business relationship is properly interwoven with the theoretical logic of the developmental state in this book, it is evident that we can shed light on the changing nature of the state's role and involvement in the economy in newly industrializing countries by paying attention to the challenging issues raised in this book. Flexibility of the state and large businesses and the relations between them, as shown in this book, can provide us with a new viewpoint for exploring the future paths of economic development in newly industrializing countries.

Hungary's Negotiated Revolution: Economic Reform, Social Change, and Political Succession, 1957-1990. By Rudolf L. Tökés. New York: Cambridge University Press, 1996. Pp. ix+554. \$64.95 (cloth); \$24.95 (paper).

Camille M. Busette
University of Texas at Austin

When an author argues that certain antecedent or structural conditions caused a particular outcome, she or he must address the Achilles heel of such types of analysis: Why, if the conditions existed all along, did the event not happen at some time other than when it occurred? Thus, with Rudolf Tökés's opening assertion that "Hungary's journey from totalitarian dictatorship to parliamentary democracy commenced on October 23, 1956" (p. 1), one anticipates that Tökés will, indeed, explain why Hungarian Socialist Workers' Party (HSWP) leaders chose to negotiate a transition to democracy in June of 1989 rather than in 1979 or in 1997. Unfortunately, this is not forthcoming. What Tökés offers is, instead, an extremely well researched compilation of political and economic trends and events in post-1956 Hungary. Essentially, his analysis of the timing of Hungary's negotiated revolution is an accretion of evidence with which he attempts to argue by suggestion that poor economic performance, sclerotic leadership and bureaucracy, and the increasing importance of young reform economists led to the HSWP's leadership's fateful decision in June of 1989 to negotiate with the opposition roundtable.

What one hopes for is a discussion of the mechanisms that transform fairly banal situations such as poor economic performance or the existence of a small and impotent community of dissent into something which forces leaders of the HSWP to change strategies. Nowhere is the elucidation of such a mechanism more necessary to Tökés's analysis than when he describes the rather meteoric increase in the power and freedom that reform intellectuals wielded in the mid-to-late 1980s. We know that in 1956 most such people were either suppressed, killed, or left the country. This would imply that dissent from 1956 on was quite insignificant. Yet, we find out from Tökés that in 1985, proto-opposition groups held a conference in the village of Monor (p. 313) and that by 1987, reform economists had published a damaging report on the economy. These facts would suggest that something transformed a negligible community of dissenters into a potent interlocutor for the communist regime. Yet, we never find out from Tökés what that something is.

To me, this omission says very much about the conceptual framework with which Tökés is working. That repression had lightened up in the 1980s enough that reform economists and other sympathizers were actually articulating alternatives to the received wisdom of the HSWP implies a startling change in the strategies that leaders within an authoritarian regime typically utilize. That Tökés fails to recognize this change suggests he assumes that, historically, dissenters were relatively politically autonomous. Such an assumption is neither consonant with the facts nor with

Tökés's own experience. Authoritarian regimes thrive on the suppression of alternative ideas, which means that citizens of those regimes will not be able to take the initiative in articulating alternatives unless the regime somehow allows them to do so. This idea implies that the articulation of alternatives that occurred in the late 1980s was a response to some changes in the strategies of HSWP leaders. In the end, Tökés's failure to focus on the mechanisms of political change in Hungary exposes the murky conceptual quality of his effort.

Some of Tökés's foci seem out of place and somewhat misguided. For example, he presents an extremely strong chapter on social change in Hungary after 1956 that appears marginal to the basic thrust of the book (pp. 117–63). His main conclusion in this section is that, yes, socialism did change Hungary's social fabric: Hungarians were more educated in the late 1980s than they had been in 1956. However, it is not at all clear how this observation advances the overall argument that economic failures, a bumbling bureaucracy, and a sclerotic leader eventuated in parliamentary democracy.

In addition, Tökés places a great deal of emphasis on the generational conflict that emerged in the late 1980s between an aging János Kádár and younger Politburo members. This emphasis seems theoretically misguided. It is important to note that János Kádár initiated a policy of political reform in 1985 and that this policy brought the HSWP to one of the critical dilemmas of its existence: How could HSWP leaders control liberalization and yet reap the benefits of the increased coalitional support it afforded? Significantly, Kádár's failure to resolve this dilemma catalyzed the intense jockeying for position in 1986–88 that Tökés so carefully documents. In sum, an overarching theoretical framework would have done much to clarify the weights of various explanatory variables.

What readers will find attractive about *Hungary's Negotiated Revolution* is its encyclopedic quality. There is no dearth of evidence here; the book is extremely well researched and represents a tremendous archival effort on the author's part. In that sense, Tökés has provided us with a definitive account of the period. More important, this book is a first-rate attempt by one of the members of the 1956 exodus to come to terms with the painful and traumatic experience of October 1956 and to recover his own place in the history of the last 30 years. It is in this sense that this work is, ultimately, a truly laudable effort.

Desert Capitalism: Maquiladoras in North America's Western Industrial Corridor. By Kathryn Kopinak. Tucson: University of Arizona Press, 1996. Pp. xvi+232. \$45.00.

Leslie Sklair

London School of Economics and Political Science

Kathryn Kopinak's study of auto-parts factories, their workers, managers, technology, and organization in Sonora, Mexico, is a valuable addition to the already substantial literature on the *maquiladora* (hereafter *maquila*) industry along the Mexico-U.S. border. The main argument of the book is that the predictions of the dual technology thesis (roughly, the change from Fordism to flexible specialization) that dominates much of the debate on industrial restructuring are not borne out in this particular instance. These predictions are that the proportions of technicians and "high-tech" operations will rise, that wages will increase, that more men will be employed in *maquila*-like plants, and that industrial districts will emerge. Kopinak makes her case with an abundance of empirical evidence and a careful analysis of the dual technology thesis and its implications. None of these predictions are confirmed in Sonora, and there is evidence to suggest that this conclusion is also valid for other *maquila* sites in Mexico (and elsewhere, though Kopinak does not comment on *maquila*-like developments outside Mexico).

Kopinak reviews the literatures on the growth and structures of the *maquila* industry, the specificities of "the North American Western industrial corridor" (Nogales, Arizona, is the third largest landlocked port of entry in the United States), and the issue of industrial restructuring with commendable economy of style. The table on page 21, "Key Dimensions in Comparison of Paradigms of Industrial Competitiveness," comparing the old paradigm of Fordism with the new paradigm of flexible specialization on the criteria of production facilities, location, work organization, and labor process is as succinct a representation of a series of complex and confusing debates as one might wish for. Methodological questions, notably why the auto-parts industry was selected for study, and the logic of the multiplicity of types of data collected (interviews with managers and workers inside and outside the factories, analysis of job advertisements, and household surveys) are dealt with quite satisfactorily. The main findings of the study indicate that "Nogales-area transport-equipment plants do not fit either alternative of the dual technology thesis" (p. 181). This is connected with the argument of Mexican researchers, notably Jorge Carrillo, whose work in Spanish, fully utilized by Kopinak, deserves to be better known, that the *maquila* industry is more adequately conceptualized in terms of "technological heterogeneity."

The strengths of the book lie mainly in the analysis of the fieldwork. For example, Kopinak makes excellent use of job advertisements in the local newspaper and interviews with *maquila* personnel in her chapter titled "Constructing the Nogales Labor Market," which permit her to

draw firm conclusions about gender segregation and technical upgrading. One response on the latter makes the point very directly:

Author: "What has your experience been with the implementation of automated technology in Mexico?"

Maquila association leader: "You would only do it when it's cost effective. As long as you could do something more cheaply by hand, why would you go out and buy a million-dollar machine?" (p. 119)

However, there are two areas where the book is less convincing. First, Kopinak seems locked into the dual technology thesis that she so effectively demolishes for her case study and leaves us wondering whether she might have been better advised to reconceptualize a new model building on Carrillo's "technological heterogeneity." Her scattered comments on how the *maquila* industry operates to insert Mexico into a more global-capitalist system of production and consumption point the way to one potentially fruitful development of her thesis. Second, the discussion of the intriguing concept of the title, "desert capitalism," is rather feeble. Her concluding chapter fails to demonstrate in any systematic fashion either that desert capitalism is a fruitful concept to encapsulate the structures and dynamics of the *maquila* industry (in Sonora or elsewhere) or that it can help to locate the *maquila* industry in the more general argument about the social and political effects of industrial restructuring.

Despite these reservations, the book is impressive and is a very capable case study of a phenomenon (export-oriented industrialization) growing in importance all over the world and why the *maquila* workers in Sonora, despite working long hours, struggle to make a decent living. The "maquilization effect"—feminizing the labor force, segmenting skill categories with most workers unskilled, lowering real wages, and getting rid of unions—is not restricted to the Mexico-U.S. border.

Ending a Career in the Auto Industry: "30 and Out." By Melissa A. Hardy, Lawrence Hazelrigg, and Jill Quadagno. New York: Plenum Press, 1996. Pp. xiii+272. \$39.50.

Angela M. O'Rand
Duke University

In 1964, General Motors (GM) and the United Auto Workers (UAW) added early retirement to their employment pension provisions, encouraging retirement with reduced benefits as early as age 55 depending on a mix of age and service criteria. By the 1970s early benefit packages became more widespread, especially within closed employment systems, and had diffused beyond the automobile industry across other sectors. This sequence of events in the United States was identified by Richard Barfield and James Morgan in their economic study of the "retirement decision" (*Early Retirement: The Decision and the Experience* [Institute

of Social Research, 1969)], which analyzed a nationally representative sample and a sample of autoworkers in the mid 1960s.

Hardy, Hazelrigg, and Quadagno contribute to this tradition of inquiry in their case study of 1,764 male GM workers drawn from an estimated 29,000 workers employed in 11 automotive manufacturing plants scheduled for closure sometime during the period 1987–90. The 1987–89 contract between GM and the UAW offered two early retirement plans: regular early retirement (RERP), and special early retirement (SERP). These plans and the reasons for their adoption or rejection by workers facing variable risks for plant closures are the principal concerns of this study.

Defined benefit plans, such as those covered by the GM-UAW contracts and accords, are seniority-based, backloaded instruments that reward long service. The benefits are calculable from formulas linking years of service to basic lifetime (monthly) benefit rates. The “30 and out” designation in the book’s title refers to the automatic maximum monthly benefit offered only to workers with the highest level of seniority under the normal and early retirement plans. This maximum benefit had operated since the Barfield and Morgan era as a strong incentive to younger workers to stay at the job until achieving maximum benefit eligibility, which then becomes an incentive to retire. A related norm, “the rule of 85,” characterizes early retirement—workers with 30 or more years of employment became eligible at age 55. Regular early retirement between ages 55 and 65 years old carried with it a normal lifetime component and a supplemental component payable until age 65.

Hence, the principal incentive structure for early retirement identified by Barfield and Morgan in the 1960s survived in its basic structure until the mid 1980s, when plant closings were a response to changing global competition in the automobile industry. Accordingly, workers ages 50 or more with only 10 years of service but facing layoffs were offered SERPs—packages with lifetime and supplemental components comparable to RERPs plus temporary benefit compensations to encourage early retirement.

Hardy, Hazelrigg, and Quadagno are interested in identifying what factors influenced workers’ responses to these offers between 1987 and 1989. They find that while 35% of their sample did not retire over the three-year contract period, 36% and 29% accepted the RERP and SERP plans, respectively. They begin with an actuarial model including measures of pension wealth (a present discounted asset value of earnings, private pension, and Social Security income streams), age, service years, and wage rate to predict early retirement. Their analysis reveals the countervailing interplay between age and seniority in response to these incentives. Retirement ages were younger among workers who left with 30 or more years of service. However, the direct effect of seniority was negative at all ages and at all service levels.

Their follow-up analyses and the final interpretation of their results are founded on their distinction between “actuarial and practical rationality.” On the one hand, these workers’ decisions appeared to reflect “an

appreciation of actuarial rationality." On the other, nearly two out of five workers who had passed the 30-year service threshold postponed their retirement. Sociologically relevant factors appeared to constrain actuarial predictions, especially among RERP adoptees who were less directly threatened by temporary or permanent shutdowns and less sensitive to pension wealth. The addition of sociologically relevant variables that represented skill levels, household income, marital status, education, race/ethnicity, overtime work, and experiences with layoffs or plant closings suggested several nonactuarial explanations for the observed diversity. High seniority was associated with a high degree of inertia, following Hirschman's terminology (*Exit, Voice, and Loyalty* [Harvard University Press, 1970]). More highly skilled workers, who incurred lower authority costs by staying, postponed retirement. Older workers with higher seniority also faced different opportunity structures for job mobility than younger workers of equal seniority and were thus more likely to postpone retirement.

The effects of contextual factors, including discussions among coworkers, union representatives, and managers, worker solidarity, and perceived generational equity were explored but with uneven results. The analysis does not disentangle the causal ordering between workplace discussions and worker decisions. The incidence and the content of workplace discussion had no relationships to the decision. Similarly, the attitudes of older workers toward the perceived pressure from younger workers that they retire yielded no systematic tendencies. A subanalysis of "atypical retirees" attempts to tease out the operation of "redistributive ethics" but reveals results in the opposite direction.

Finally, the analysis of postretirement satisfaction reveals nothing new. Retirees tend to be satisfied with their retirement, and self-selected male retirees drawn from a more pension-rich environment are even more likely to be satisfied.

All in all, this book adds to the retirement literature by replicating the strong effects of pension provisions on work exits and by contextualizing this process in the 1980s when plant closings were more prevalent on the industrial landscape. It succeeds less well in specifying the social influences bearing upon retirement decisions above and beyond pension incentives and immediate work-related opportunity structures. It also bears little relevance to the predominant pension landscape of the more services-dominated workforce of the 1990s and beyond—which includes a significant share of women workers and which is less likely to be covered by pensions, especially by the defined benefit pensions that constituted the framework of pension incentives between 1960 and the 1980s in closed employment systems.

Living Rooms as Factories: Class, Gender, and the Satellite Factory System in Taiwan. By Ping-Chun Hsiung. Philadelphia: Temple University Press, 1996. Pp. x+182. \$44.95 (cloth); \$18.95 (paper).

Kuang-Chi Chang
University of Chicago

For the past 30 years, the Taiwanese economy has grown at an average annual rate of 9.1%. This "economic miracle" is largely built on its export sector, which accounted for 54.3% of Taiwan's GNP in the 1970s and 41.4% in the 1980s. One unique feature of Taiwan's export economy is its reliance on the "satellite factory system," which provides low-technology, low-value-added, labor-intensive products for sale overseas. Most of these family-centered factories employ fewer than 30 workers, the bulk of whom are married women. The fact that these small firms often subcontract work out to housewives working from their living rooms forms the title of Ping-Chun Hsiung's study, echoing the official government slogan, "Living Rooms as Factories."

Using ethnographic data collected from several satellite factories involved in the production of wooden music boxes, Hsiung vividly portrays the everyday lives of these married working women and how, in the process of Taiwan's industrialization, these women are now expected not only to serve their families in their traditional roles as dutiful Chinese wives, mothers, and daughters-in-law, but also as industrious low-paid or unpaid foot soldiers in Taiwan's capitalist economy. These dual, often conflicting roles, she argues, are the product of two underlying mechanisms in Taiwanese society—the patriarchal mechanism of traditional Chinese society in which women provide reproductive labor to the family and the capitalist mechanism in which women provide cheap, unskilled labor to the factory.

Hsiung shows how, in the competitive international market, women's labor works well by reducing uncertainty and overcoming some structural constraints. The "living room as factory" system allows even small Taiwanese firms to reduce overhead, lower labor costs, transfer risk, and mitigate fluctuations in the demand cycle. These women are often paid less than the legal minimum wage and on a piece-rate basis, and since they are assumed to not be their families' breadwinners, they are laid off with impunity during recessions, while they provide unregulated overtime labor during peak business seasons.

Most interestingly, Hsiung argues that Taiwan's ruling Kuomintang (KMT) party has played an essential role in promoting the satellite factory system and in reconciling the conflicting roles that the patriarchal capitalist system imposes on Taiwanese women. Hsiung claims that the KMT is historically conservative and antifeminist and, consequently, a strong proponent of traditional women's roles as seen through such government programs as "Mothers' Workshops," where women are taught "feminine ethics and family values" in community classes on cooking,

sewing, flower arranging, interior decorating, and so forth. At the same time, however, the KMT's own political fate is inextricably tied to the success of the national economy. It has established an institutionally oppressive labor regime, countering the labor shortage and rising wage rates plaguing the industrialists with the "living room as factories" program. As such, the satellite factory system represents a state-instituted compromise between the conflicting social interests of capitalism and patriarchy. As Hsiung notes, "On the surface, women *choose* to work in the small satellite factories in their neighborhoods. In essence, this is the only way married women can fulfill the latest multiple expectations of a married working-class woman: a responsible mother, dutiful wife, filial daughter-in-law, and hardworking worker" (p. 101).

Using statistical data, Hsiung shows that more working men than women have achieved upward mobility by becoming self-employed workers or owners of small factories. When some women experience upward mobility by marrying the factory owners or by helping their husbands start family businesses, they separate themselves from the other women who work in the system as wage workers or home workers. Although women from both groups recognize existing gender inequality, their shared gender consciousness does not transcend class boundaries. Compared to the interests that the owners' wives share with their husbands, the bonds between them and other female wage workers seem trivial and offer no barrier to potential conflict. Hsiung also offers interesting insights on how, in the face of state-sponsored controls on organized labor, workers at repressive satellite factories employ systematic though informal and even clandestine methods to protest unfair treatment.

Sociologists interested in gender studies, family, social stratification, labor issues, political economy, and economic development will find this book of great value. It is well organized and provides many new insights into these areas. Hsiung has achieved a lot in this book, more than its title would indicate. She calls attention to an important area overlooked by other sociologists. By doing so, she offers a new angle on how Taiwan created its economic miracle. While many scholars studying Taiwan's economic growth have focused on macrolevel factors such as state industrial and labor policies, Hsiung builds a link between micro- and macrolevel issues by discussing the individuals living under these policies and how these macrolevel factors affect people's day-to-day experiences, choices, and sacrifices, and how these choices and sacrifices in turn advance the macrolevel policies.

Nonetheless, it may be argued that Hsiung places too much emphasis on the structural forces of capitalism and social patriarchy, while leaving out other important issues. She implies that the state policies embodying patriarchy and capitalism function as a "pushing" factor influencing all Taiwanese women. Yet some groups of women choose to work in the satellite factories while other groups of women do not. Other factors such as family income, socioeconomic status, education or professional train-

ing, and past work experience suggest that government policies act more as a "pulling" force that draws from an existing pool of low-income, low-skilled, and uneducated women at a particular time in a nation's industrial development. When Taiwan's economy was first taking off and per capita GDP was very low, many Taiwanese families passed through financial difficulties by converting their living rooms into factories. Today, many of these women have grown out of the need for this supplemental income as their families (and Taiwan) have grown out of poverty.

As societies become richer, problems like labor shortages and rising labor costs emerge. In recent years, many Taiwanese satellite factories have either moved to other newly industrializing countries such as Indonesia or China or have closed down. How Taiwan's satellite factory system and its women workers are influenced by this situation would be an interesting topic for future research.

Heroic Defeats: The Politics of Job Loss. By Miriam A. Golden. New York: Cambridge University Press, 1997. Pp. xvi+195. \$59.95 (cloth); \$18.95 (paper).

Chris Tilly
University of Massachusetts at Lowell

There are two reasons to read *Heroic Defeats*. First, read it if you are interested in why and how strikes take place—particularly strikes over job loss. Miriam Golden attacks these questions by way of analytical comparative case studies of five major labor conflicts in postwar Britain, Japan, and Italy. Second, read the book if you care about debates over general models of social action. Golden uses her empirical study to underpin a forceful argument in favor of rational actor models.

Golden starts by asking, Faced by large-scale layoffs, why do unions at some times react with strikes and at other times not? In response, she advances two central propositions. Unions only strongly resist job loss when it poses a serious threat to the union organization itself—particularly by threatening the discharge of a substantial number of shop floor activists. And despite appearances, strikes are rational undertakings by companies and unions, not outbursts fueled by anger or zeal.

Stating that strikes are rational means that both parties make the best calculation they can of the likely costs and benefits of different courses of industrial action and then choose the action with the greatest payoff. Since the costs of a strike exceed the benefits for at least one party (i.e., somebody loses every strike), strikes should only happen under one of two circumstances: either incomplete information prevents at least one party from correctly predicting the outcome of their action or the parties must receive subsidies from third parties to offset their losses.

Am I starting to sound like an economist? It is not accidental. Golden, a political scientist, uses the microeconomic tools of game theory to con-

struct her model. But less technically inclined readers need not fret. The model is technically simple, its implications intuitively clear. And by far the bulk of the book consists of the textured case studies the author uses to test her model. In this enterprise, Golden claims success for the model. Unions struck over job loss in the Japanese coal mines in 1953 and 1959, at Fiat in 1980, and in the British mines in 1984 because layoffs targeted a critical number of union activists. Unions did not strike over mass layoffs at British Leyland automobiles in 1979–80 nor in the United States for most of the postwar period because entrenched seniority provisions protected most activists (who tend to be senior workers) from dismissal. Fiat management sparked the 1980 strike by misestimating the point at which the union would react to layoffs. British and Japanese coal mine managers and unions entered battle buoyed by outside assistance—on the one side from employer associations and conservative governments, on the other side from union confederations.

Golden's carefully documented empirical conclusions are for the most part quite persuasive. There is something of a gap, however, between the tidy models in the second chapter and the dense analysis that explains the strikes themselves. Consider five characteristics of the models: (1) they only address one-sided incomplete information (the employer does not know how strike-prone the union is); (2) the management layoff decision is just a yes/no decision, not a decision about the number of workers to lay off; (3) labor and management can each be summarized with a single set of preferences; (4) unions strike to prevent their activists from being fired; (5) the strike is a single move in the game, with a determinate outcome. In the actual labor disputes documented by Golden, exceptions to these simplifications play important parts in shaping the outcomes. For example, the British miners' union was taken by surprise by the harsh police response to picketers; Fiat's unions accepted small layoffs in 1979 but struck in response to a larger number in 1980. The formation of alternative antistrike unions in midstrike helped doom Japanese and British coal strikes, while the triumph of a hawkish wing of management torpedoed a settlement in the Japanese case; Fiat unionists seem to have struck despite foreseeing defeat—perhaps because an outcome of layoffs accompanied by protests would do more to strengthen union organization in the long run than the same layoffs met with acquiescence.

In short, labor disputes are messy affairs. Each party typically has only a vague notion of the strength of its own forces and its adversary's and the resources that third parties are likely to provide. Each pursues multiple objectives and wrestles with internal divisions. Strategies consist of step-by-step responses to often unanticipated situations. This does not mean that the actors are not rational, but it does mean that their rationality is sharply bounded, their objectives plural, their choices contingent—all of which distances them from the standard rational actor model.

All the more impressive then, if theory can draw out nonobvious regularities from this tangled reality. Golden's strongest candidate in this regard is her notion that unions strike over job loss to defend activists from

being fired, not—despite their public positions—to avert the job loss itself. The evidence she presents, in particular the contrast of Fiat with British Leyland, while not overwhelming, is compelling. This insight alone makes *Heroic Defeats* well worth reading.

When Work Disappears: The World of the New Urban Poor. By William Julius Wilson. New York: Alfred A. Knopf, 1996. Pp. xxiii+322. \$26.00.

Joe R. Feagin
University of Florida

In his now forgotten first book, *Power, Racism, and Privilege* (Free Press, 1973), William Julius Wilson argued that "racial stratification is largely a function of dominant-group [white] efforts to control scarce resources either by eliminating or neutralizing subordinate racial members as competitors or by exploiting their labor" (pp. 189–90) and that racist ideologies of whites have justified black exploitation. In *When Work Disappears*, Wilson seems to have forgotten his argument about the centrality of white-racist thought and action in creating the stratified and oppressive conditions still faced by African-Americans.

Drawing on new data and rich field studies, Wilson focuses on the impact of job disappearance on the lives and culture of Chicago's inner-city residents. A major focus is on how job-troubled environments create negative values and behaviors, such as drug pushing. The values that develop where jobs disappear are often negative and entail less connectedness with the work ethic. Mixing statistical data with rich quotes from the field, Wilson offers a vivid portrait of conditions for African-Americans living in areas of high unemployment.

Wilson's approach to the shaping of the values of the poor is structural. He underscores the role of broad economic changes, including suburbanization and internationalization of jobs, in creating inner-city problems. He accents the discrepancy between the inner-city location of the poor and the suburban location of many new jobs being created. However, this argument has been questioned in other recent research, which has found that black workers who do travel to look for suburban jobs—often from homes in black suburbs—face much job discrimination.

Unlike his previous two books, Wilson's new book does briefly assess discrimination's role in inner-city problems, and he critiques conservatives who accent racial-genetic explanations for poverty. Nonetheless, he does not offer a substantial discussion of discrimination as a major barrier for black urbanites and seems to reject the view that racism is a deep-seated problem. Only one section, on employer attitudes, significantly discusses racial matters. Wilson reports finding some statistical discrimination (inner-city employers screening out individual applicants based on work images of groups). Still, he plays down discrimination by white employers, noting that few of them reported discrimination to be a problem

for inner-city black men in their interviews and that black employers were often negative in their views of inner-city employees.

A major problem is Wilson's neglect of Chicago's racial history. White politicians and business leaders have often worked to segregate black workers and communities. Chicago has a long history of housing segregation, including violence against black families moving into white areas. In terms of the index of dissimilarity, Chicago remains one of the nation's most highly segregated cities; real-estate discrimination has long been noted as a major factor determining where black citizens live. If Chicago is like other cities recently examined in housing-audit studies, 70%–80% of whites with housing to sell or rent likely discriminate against African-Americans. Widespread housing discrimination and segregation have profound effects on the educational, business, and job opportunities that are central to Wilson's account.

One critical problem for black Chicagoans, Wilson demonstrates, is the high concentration of poor with problems in certain neighborhoods. Wilson accents the import of this concentration and the movement of better-off African-Americans and black institutions from the worst inner-city neighborhoods. Yet, he does not discuss the fact that housing discrimination and segregation are significant reasons for past and present black concentration. In his first book, *Power, Racism, and Privilege*, Wilson accented the central importance of powerful whites orchestrating patterns of racial relations, and this new book would have been stronger if he had brought this theme back by assessing the role of powerful business and political elements in shaping residential and community patterns in Chicago.

Wilson concludes with a considered call for new educational programs and employment policies, including job-information programs and WPA-like programs to increase employment opportunities. His prescriptions for expanded social programs are practical and bold in the current political climate. Among major industrial nations, the United States has provided the weakest job-training and job creation programs. One important reason for this is that the federal government is substantially under the control of a corporate elite that usually seeks to weaken labor, cut taxes, and reduce government social-welfare programs. There is much data supporting Wilson's argument that international business competition has led to many firms closing plants in cities like Chicago. However, movements of jobs to suburbs and overseas are not part of impersonal or natural fluctuations in the global economy but rather signal that certain ascertainable U.S. investors are moving much capital, on a long-term basis, to countries with cheap, exploitable labor and weak state regulation of businesses. Without a radical restructuring of the highly undemocratic investment framework of capitalism, it is not possible to deal fundamentally with job disappearance problems in central cities.

The Origins of the Urban Crisis: Race and Inequality in Postwar Detroit.
By Thomas J. Sugrue. Princeton, N.J.: Princeton University Press, 1996.
Pp. xviii+375. \$35.00.

Bobby M. Wilson
University of Alabama at Birmingham

The Origins of the Urban Crisis, as the author notes, is about blacks in Detroit, which once had a secure status and was once considered a promised land for blacks, but that found itself increasingly marginalized due to economic restructuring. Like so many other urban destinations for blacks migrating north in the era between the Great Depression and World War II, the promised land was just a promise and not a reality. At the time of the Great Migration, the automobile industry, which attracted blacks to the Motor City, was beginning to downsize and relocate its plants out of the central city to rural and suburban locations.

From this book we learn that a mobile and flexible capital is not a phenomenon of late-20th-century capitalism but is very much a phenomenon of mid-20th-century capitalism. As Sugrue noted, "The rusting of the Rust Belt began neither with the much-touted stagflation and oil crisis of the 1970s, nor with the rise of global economic competition and the influx of car or steel imports. It began, unheralded, in the 1950s" (p. 6). The mobility and flexibility of capital has since increased to produce the age of global capitalism. But the forces of deindustrialization affected blacks in Detroit and other cities in the American manufacturing belt much earlier.

Deindustrialization served to boost unemployment in Detroit's black community, wreaking havoc on and trapping blacks in the inner city. Neoclassical economic proponents assume that as capital moves so do workers who can move without social constraints to areas of new economic growth. Sugrue clearly shows that this is not the case. In Detroit, racial discrimination in both the workplace and in residential areas prevented blacks from relocating as capital moved from the inner city to the suburbs. Neither civil rights nor the welfare state could counter the increasing mobility and flexibility of capital that produced deindustrialization and overmanned spaces in the inner city of Detroit. Labor unions also proved unresponsive. Deindustrialization wreaked havoc on organized labor, which was never truly organized across racial lines.

America has never produced conditions that foster class identification as a category with effects on everyday existence. Labor and urban political machines carefully cultivated race and ethnic identities of immigrant workers. According to Sugrue, "Assumptions about racial difference were nourished by a newly assertive whiteness, born of the ardent desire of the 'not-yet-white ethnics' to move into the American mainstream. To be fully American was to be white" (p. 9). This assertiveness was not new. These ethnics have been constructing a notion of whiteness ever since they crossed the ocean to America in the late 1800s. Residential

segregation would provide the staying power for racial differences. As a result, "unions often reinforced or quietly acquiesced in employers' discriminatory hiring and upgrading policies, and only seldom challenged management decisions on plant location and expansion that had significant long-term ramifications for Detroit's workers" (p. 11). Thus, the blame for a lack of working-class response is to be found within the working class itself: the inability of American workers to formulate a working-class culture and identity that transcend race, gender, and ethnicity.

The Origin of the Urban Crisis is mainly a description of the economic and social forces that produced the crisis in Detroit during the postwar era. Like most descriptive work, one may find it difficult to keep an account of all the places, events, and actors that Sugrue considered critical in his description of the postwar crisis in Detroit. Providing a theoretical framework for the economic restructuring that produced the crisis would have been enriching. I am surprised that he did not consider "regulation" theory, (e.g., M. Aglietta, *A Theory of Capitalist Regulation: The U.S. Experience* [New Left Books, 1979]), which is widely used to analyze the very restructuring of the economy that produced the crisis in Detroit. In the introduction the author writes, "The American economy . . . automated production and relocated plants in suburban and rural areas, and increasingly in the low-wage labor markets of underdeveloped regions like the American South and the Caribbean" (p. 6). This is clearly a tendency of the regime of accumulation that regulationists call "Fordist." The Fordist regime extended successful industry and its firms into new territory, thereby creating what A. Lipietz called "peripheral Fordism" ("New Tendencies in the International Division of Labour: Regimes of Accumulation and Modes of Regulation," in A. Scott and M. Storper, eds., *Production, Work, Territory: The Geographical Anatomy of Industrial Capitalism* [Allen and Unwin, 1986], pp. 16–40). Regulationists, however, have not focused much on the role of race in regulating the economy. Sugrue provides a wealth of information about race and economic restructuring for regulationists to ponder. It is a book that should be read by all who are concerned with the impact of restructuring on local communities.

It Takes a Nation: A New Agenda for Fighting Poverty. By Rebecca M. Blank. Princeton, N.J.: Princeton University Press, 1997. Pp. xvii+340. \$29.95.

Nancy A. Denton
State University of New York at Albany

From the first to last page of *It Takes a Nation*, the reader is in for a treat. Written for a nontechnical audience by a prominent economist, the book summarizes past and current poverty research and offers expert opinion on future policy options. Encyclopedic in coverage yet intensely

readable, the book is a treasure because all the information about poverty programs is in one place. By using questions as headings, Blank tunes into her readers' needs. By cleanly abstracting and interpreting poverty research, she provides the opportunity for theoretical reflection on poverty in its entirety, offering the potential for new insights by correcting our consistent misunderstanding of the nature of poverty and the poor.

The first four chapters of *It Takes a Nation* emphasize poverty definition, history, and program results, while the last three chapters emphasize the author's recommendations and rationale for policy. The introduction offers three "lessons," which are then fleshed out in the first three chapters: (1) that the belief that poverty is more behavioral, more ghetto based, and more a problem experienced by people of color is an error, (2) that jobs are less effective at reducing poverty since the primary change in the lives of the poor over the past 20 years is the deteriorating set of economic opportunities available to less-skilled workers, and (3) that the nihilistic reaction of "nothing works" is a serious misinterpretation of history.

Chapter 4 evaluates the effectiveness of poverty programs and simultaneously instructs the reader in how to scientifically evaluate them by looking at the full range of program effects and weighing them, for a program may have an undesirable effect, but the cost of abolishing it may be higher still. For example, cutting welfare aid to unwed mothers punishes the mother but what is the cost of letting her children grow up in extreme poverty? Blank also says that we should not evaluate all programs by the poverty rate, but rather use the correct yardsticks: Do nutrition programs raise nutrition? Does health insurance improve health? Can states that pay a maximum of \$3,000 in Aid to Dependent Children get families above an \$11,900 poverty line? Blank asks us to not evaluate programs by looking only at the behavior of the participants: in hard economic times, a jobs program may keep participants' unemployment stable while nonparticipants lose jobs. Participants' unemployment may only decline in good economic times.

Equally compelling are the author's policy recommendations. Chapter 5 points out that most private charities receive income from the government. Since current government spending is about 20 times that of charities, it is unreasonable to think they can make up the gap. Chapter 6 outlines criteria for effective targeted programs, which will be more expensive because one must monitor behavior as well as hand out money. Even effective programs are potentially problematic because savings will not accrue to the targeted programs but to other budgets (if adolescent programs reduce crime, the police budget saves money) and because they will not necessarily affect the poverty rates caused by the deterioration of labor market opportunities for less-skilled workers. In the final chapter, a three-tier family assistance program is proposed, focused on the question, What do you need to get back on your feet? Persons would first be given short-term help, then job training, with cash assistance only as a last resort, all tiers to be supplemented with expansions of the EITC, Child

Care Tax Credits, and Assured Child Support, as well as a link between federal monies for poverty and the economic cycle since poverty is countercyclical. All policy recommendations emphasize family, work, and the recognition of specific needs of different groups among the poor, allowing us to see how a "one size fits all" poverty policy will no longer suffice.

Any reader of a book this wide ranging is sure to find pet topics slighted, and I am no exception. One image the book is trying to convey is the similarity between the poor and the rest of us. Thus I was disappointed that the broad-based \$54 billion middle- and upper-class welfare program in the form of tax deductions for the interest on home mortgages and local property taxes was never mentioned (more than double the housing assistance for the poor or Aid to Families with Dependent Children, p. 86). Likewise, issues surrounding poverty definition were too briefly discussed, missing the chance to point out that people now spend about one-fifth of their incomes on food, not one-third as is assumed by the current poverty definition or that the relationship between the poverty level/minimum wage and the median income has been eroding steadily over time. Though minor, these omissions reinforce the idea that the nonpoor do things on their own without government assistance.

In closing, I was left with an overwhelming sense of the balance of the presentation and the author's own positions on poverty. If this book were required reading before *anyone* was allowed to make a public statement about or enact a program on poverty, the nation would have taken an enormous step toward fighting poverty.

Saving Our Children from Poverty: What the United States Can Learn from France. By Barbara R. Bergmann. New York: Russell Sage Foundation, 1996. Pp. xiv+184. \$34.95.

Suzanne M. Bianchi
University of Maryland

In *Saving Our Children from Poverty*, Barbara Bergmann examines the French welfare state and uses what she finds to outline a system of support for children and their families that might be efficacious in alleviating child poverty in the United States. According to Bergmann, it is a book written for the future when we "have a president who can effectively frame and forward the required agenda, a time when generosity toward the 'have nots' is greater, and when antigovernment rhetoric has been overcome by an even more obvious need for action" (p. 151).

After a brief introduction highlighting similarities between France and the United States in female labor force participation, births outside marriage, and rates of economic growth and investment (also noting the much larger population of the United States and the higher rates of unemployment and taxation in France), Bergmann catalogs French spending on programs for children. She shows that on a per capita basis the French

spend about 60% more on child well-being than the United States does—with beneficial outcomes: much lower child poverty rates, lower infant mortality, more incentive for parents to be employed, and less stigma attached to receipt of government assistance.

Much of the book is devoted to a detailed description of the complex array of French programs that benefit children and their families—programs providing education and child care, enhancing family income, and ensuring the health of children and their families. As Bergmann explains and the reader will no doubt agree, the complexity and number of programs in France make it a system that the United States would not want to duplicate. However, the results are worthy of emulation—and Bergmann seeks to achieve these with her own program of “help for working parents,” outlined in chapter 7 and advocated in chapter 8.

France has an extensive system of high-quality universal early education/child care that begins when children are two and a half years old—a system that is conducive to parental employment. France also has universal health insurance and an active intervention program that facilitates healthy birth outcomes, child immunizations, and so on. France also has a number of income support programs for low-income families, such as assistance with housing, that augment earnings. Bergmann’s plan for ending child poverty in the United States would include vouchers for quality child care, health insurance for all children and their families, and expansion of existing programs, such as tax credits and food stamps, in order to bolster the family income of those employed in minimum wage jobs.

The book makes a compelling argument, and the description of American programs and alternatives is superb. What is missing is a full discussion of the assumptions that undergird Bergmann’s proposed program. First, the program assumes that it is good for mothers to work for pay—for adult women and men to spend their time similarly rather than differently. Second, it assumes that young children will be better-off in high-quality day care settings than they will be with their own mothers, especially poor unemployed mothers. Third, it assumes that the most effective, and also most politically palatable, programs for the support of poor children are ones where choices are made for poor parents—programs that do not give cash but rather use vouchers to ensure that parents do not forgo high quality care for their children or health care coverage for themselves and their children. Except for a discussion of vouchers versus cash—a point Bergmann feels she must address because fellow economists will be uneasy with her “constrained choice” approach—there is no elaboration of these underlying assumptions.

The reader is left with the following questions: What is there to compel the American public and policy makers to accept the vision of gender equality that is needed for Bergmann’s program—a vision that asserts that it is in women’s short- and long-term interest to continuously engage in paid work, much as men do, and that accepts that what is good for mothers (i.e., allocating time to paid work even when children are young)

is not detrimental to child well-being? Is there any evidence to support the assumption that children will turn out better if they do not spend their preschool years with welfare mothers but, rather, spend most of their time in high-quality day care settings? Further, what happens to children who spend their preschool years in nurturing day care settings but then get coughed back into the same low-quality educational settings that serve the neighborhoods where their low-paid mothers (and fathers) can afford to live? Finally, is Bergmann right that poor parents—even middle-class parents—often do not make decisions in the best interest of their children or, at least, that the American taxpayer is not convinced that the poor can make sound decisions? *Saving Our Children* presents the beginning of a persuasive argument about what to do for poor children and why. Fully specifying and supporting the underlying vision is the book still waiting to be written.

Development, Change, and Gender in Cairo: A View from the Household. Edited by Diane Singerman and Homa Hoodfar. Bloomington: Indiana University Press, 1996. Pp. xl+195. \$39.95 (cloth); \$17.95 (paper).

Valerie J. Hoffman
University of Illinois

The seven essays in this book constitute a major challenge to the standard exclusion of the study of households and women from the analyses of political economy, as well as a challenge to the very notion of a strict dichotomy between public and private spheres in Muslim societies. Rich in data, based on significant extended fieldwork in the communities they describe, *Development, Change, and Gender in Cairo* documents the manner in which lower-class Egyptian households function as economic and political units and the manner and extent to which they interact with the apparatus of the state. Throughout the book we obtain a glimpse of how the society abides by and exploits its "familial ethos," and we gain an appreciation of the amazing courage, industry, and ingenuity of ordinary men and women to seek advantages for themselves against overwhelming odds—how people we ordinarily see as "victims" become shrewd managers of their own affairs. Certain common practices that are often interpreted negatively are depicted here as rational strategies: for example, when a relatively poor family purchases major electrical appliances, this is sometimes criticized by outsiders as inappropriate and conspicuous consumption, but Hoodfar explains this as a strategy of economic investment. A society that does not participate in the official banking system devises other strategies for the saving and accumulation of capital, such as founding a neighborhood cooperative society or investing in electrical appliances that can later be sold when capital is needed, an apparently effective strategy in lower-class Cairo, where the secondhand market is thriving. In short, we see a society where informal

networking structures are often more important than the official structures that preoccupy the analyses of political scientists.

The importance of microanalysis crops up repeatedly here. While the impact of Sadat's "Open-Door" economic policy has frequently been discussed by academics interested in Egypt's economic health or its role in fueling Islamist opposition, the impact of economic liberalization and spiraling inflation on the lives of ordinary people is rarely discussed. Hoodfar's analysis of the impact of large-scale male migration to work in the Gulf countries adds considerably to our understanding of this phenomenon and challenges the sweeping generalizations that are often made on this subject. K. R. Kamphoefner's analysis of women's attitudes toward literacy and education in light of the inefficacy of Egypt's educational system and its failure to prepare students for careers is far more insightful than a simple analysis of statistics. Unfortunately, her contention that the educational system has failed to cope with gender differences is left somewhat unexplained, and no suggestions for gender-based modification are made nor does she acknowledge the dangers implicit in the suggestion that education ought to take account of gender differences. Nawal Hassan's contribution is different from the others in the book—concise and programmatic rather than analytical, reflecting the interest and long-term involvement of this extraordinary woman in the welfare of the traditional urban quarters of Cairo. Her article is an intriguing glimpse into the agonized and often irrational workings of Egyptian bureaucracy in response to a housing crisis and the need to preserve historic monuments in the 1970s. The other essays rely on research done in the late 1980s and early 1990s.

The focus on households necessarily provides a new focus on women. The enormous female participation in economic production, including income-producing activities, is described here. Singerman, whose essay concludes the volume and is the most comprehensive of all, writes, "When one ventures inside of Egyptian homes, stereotypical images of the patriarchal family fade. Within the home women are active, demanding participants. . . . Parents are not obeyed unquestioningly by children, nor husbands by wives" (p. 152). On the other hand, the essays by Hoodfar, Arlene MacLeod, and Kamphoefner all suggest that work outside the home is not necessarily empowering for women. The editors are aware that their findings could be exploited by advocates of traditional gender norms and stress that their goal is descriptive, not prescriptive.

Despite the editors' admission in their introduction that the authors' study of different communities sometimes led to different conclusions, the differences between the communities that were studied are not always clear. There are important differences in gender relations between families of the old urban areas of Cairo and immigrants from the Delta, on the one hand, and in Upper Egypt, on the other. Women from traditional urban areas of Cairo and the Delta have acquired a reputation for personal assertiveness that borders on the legendary. Upper Egyptian society is much more frankly patriarchal, and its women are more restrained and

submissive. A lack of specificity about the population studied can obscure differences among lower-class Cairenes: Kamphoefner, for example, describes her population sample as *baladi*, although this term applies only to traditional urbanites, not the immigrants she describes. Nonetheless, the quality of each of these essays is excellent, and the book warrants extensive reading by political scientists, sociologists, and all scholars of the contemporary Middle East.

Latinos in New York: Communities in Transition. Edited by Gabriel Haslip-Viera and Sherrie L. Bayer. Notre Dame, Ind.: University of Notre Dame Press, 1996. Pp. xxii+338. \$34.95.

Héctor Cordero-Guzmán
New School for Social Research

In *Latinos in New York: Communities in Transition*, Gabriel Haslip-Viera and Sherrie L. Bayer contribute to the study of race and ethnic relations and to the growing field of Latino studies. The essays included in this volume also provide important insights on the nature and extent of recent social change in New York City and add to this voluminous literature as well.

Size and demographic composition make New York City a unique and central place for the study of the Latino and Afro-Caribbean experiences in the United States. The essays included in this volume explore the increasing relevance and complexity of the Latino community by showcasing recent material produced by members of the faculty at the City University of New York (CUNY) and by researchers who have been affiliated with that institution. The topics discussed in the book are divided into two main sections: "Historical and Sociological Perspectives" and "Policy Issues."

The first section includes a well-documented chapter on changes in the size and composition of the Latino population in New York City from circa 1820 to 1990 (by Haslip-Viera); a detailed analysis of socioeconomic conditions in the Dominican community and the relationship between economic underdevelopment in the Dominican Republic and the economic segmentation and marginalization of the community in New York City (by Hernandez and Torres-Sailant), a chapter on the growing Mexican population and its incorporation into New York City's racially segmented socioeconomic structure (by Smith), two chapters on the contemporary meanings, uses, and impacts of "race" and "identity" on Latinos (by Rodriguez and Hernandez), and a history of Puerto Rican marginalization within the Catholic Church (by Diaz-Stevens). The second section on "Policy Issues" begins with a chapter on economic segmentation and wage inequality (by Melendez), an excellent essay on community struggles around access, resources, pedagogy, and curriculum in schools and universities (by Vazquez), a theoretical discussion of approaches to Puerto

Rican politics focusing on how the concentration of "social power" prevents full political participation by the community (by Sanchez), and a final chapter on the impact of the 1986 Immigration Reform and Control Act (IRCA) on New York City's Latino community (by Bayer).

Most of the chapters in this book are nicely written and well documented. However, the volume has a number of limitations that compromise its usefulness as a comprehensive analysis of the Latino experience in New York City. First, the introduction has a nice description of the literature but does not engage in a sophisticated theoretical discussion of the arguments that have been proposed to explain the particular location of Latinos in New York's socioeconomic structure. The absence of a solid theoretical grounding deprives the volume of a distinct perspective and loosens the connections that are and can be made between the articles included in the collection.

Second, while stratification and inequality in the Latino community are central themes throughout the volume, the article that focuses directly on the topic is disappointing. Theoretically, the article adopts a supply and demand view of the labor market at the expense of a more sophisticated construct focusing on the political economy of migration, socioeconomic networks, employment structures, social resources, and discriminatory sorting. Empirically, both the data and the national origin categories used in the article are not the most adequate. The data comes from the 1980 census (16 years old by the time of publication) and the national origin categories used are: white, black, Mexican, Puerto Rican, Cuban, and other Hispanic. These groupings are appropriate for national comparisons but unacceptable in terms of a comprehensive analysis of the Latino experience in New York City. The Mexican and Cuban populations were relatively small in 1980 while the white, black, and, particularly, the "other Hispanic" categories (see pp. 14–15) are large and heterogeneous. New York City's complex and revealing history of economic, demographic, and fiscal change is not adequately highlighted in the volume, and the material presented does not engage well-publicized monographs and reports like New York City's Department of City Planning's *Puerto Rican New Yorkers in the 1990s* (Department of City Planning, 1994); F. Rivera-Batiz and C. Santiago's *Puerto Ricans in the United States: A Changing Reality* (National Puerto Rican Coalition, 1994); Andres Torres's *Between the Melting Pot and the Mosaic* (Temple University Press, 1995); and Orlando Rodriguez, et al.'s *Nuestra America in New York* (Hispanic Research Center, Fordham University, 1996).

My third general comment is that the conclusion, written by Juan Flores, is theoretically weak. This essay does not discuss the nature, causes, and consequences of the uneven educational, social, economic, and political processes described throughout the volume and does not draw any lessons or potential policy implications from the research.

In spite of some limitations *Latinos in New York: Communities in Transition* addresses research questions and describes sociohistorical processes that have local and national relevance. At its best, the volume

sheds light on contemporary debates over community development, migration, assimilation, stratification, and racial/ethnic inequality. This book should be considered by scholars interested in all of these topics and used to supplement the reading lists of graduate and undergraduate courses.

Life Lines: Community, Family and Assimilation among Asian Indian Immigrants. By Jean Bacon. New York: Oxford University Press, 1996. Pp. xv+295. \$49.95 (cloth); \$19.95 (paper).

Nazli Kibria
Boston University

Life Lines is a welcome addition to the currently limited sociological literature on the assimilation of Asian Indian immigrants. The book's approach toward assimilation is a social-psychological one, emphasizing the interactive and reflective mechanisms that characterize how Asian Indians situate themselves and make sense of their identity in the United States.

Life Lines is based on qualitative research conducted in the Asian Indian community in Chicago. The book begins with an analysis of community organizations, drawing on data collected through participant observation, interviews with community leaders, and content analysis of stories in the Asian Indian immigrant news media. The bulk of the book however, consists of portraits of five immigrant families. Piecing together materials from life-history interviews with family members, Bacon presents wonderfully rich and vivid narrative accounts of how family members, through processes of interaction and reflection, make sense of who they are in the context of the United States. Drawing on these materials, the book then moves on to outline "a model of intergenerational change," one that tries to capture how the process of assimilation unfolds within intergenerational contexts. A major element of the model is "parental history," which affects the ways in which family members approach issues of identity. Also noted is the "x-ness dynamic," or a sense of difference from the mainstream. For Asian Indians, this translates into the task of defining and negotiating "Indian-ness" in relation to the mainstream.

I recommend *Life Lines* to those who are interested in the topic of contemporary immigrant assimilation. It is, furthermore, absolutely essential reading for those who are concerned with South Asian immigrant communities, not only in the United States but in other parts of the world as well. Bacon's analysis of how elements of an Indian worldview shape and play out in Indian immigrant community life is extremely insightful. Also valuable are her observations about the relationship of community to assimilation. Like a number of other contemporary immigrant groups, Asian Indians in the United States have not, for the most part, clustered together in residential enclaves. They thus defy the traditional image of

tightly knit, geographically bound immigrant communities. Bacon shows us that such communities nonetheless mediate assimilation processes. They do so by providing a standard rhetoric and set of prototypes by which to define and talk about the problematics of assimilation, such as the essential differences between Indian and American cultures. These ideas and language filter into families, who use them in varied and complex ways to deal with the task of situating themselves in the United States.

Many of *Life Line*'s insights about immigrant assimilation would have been greatly enhanced by some consideration of the structural location of the families studied. As I read the book, I found myself continually thinking about how class, race, immigrant, and gender status could have been shaping the kinds of interactions and negotiations of identity so painstakingly described in the book. In a number of the family portraits, I struggled to find some clear mention of the specific jobs or occupations of family members, which might have shaped their approach toward identity in critical ways. On a related note, the inclusion of some comparative materials on assimilation processes among other groups might have helped to structurally situate these Asian Indian immigrant families. In all fairness, Bacon begins *Life Lines* by saying that she is dealing with the social-psychological rather than the structural aspects of the assimilation process. But it is unclear to me that these two dimensions can be effectively separated from each other to the extent that they are in this book.

Metropolitics: A Regional Agenda for Community and Stability. By Myron Orfield. Washington, D.C.: Brookings Institution Press, 1997. Pp. xvi+224. \$28.95.

Allen J. Scott
University of California, Los Angeles

One of the main expressions of the crushing dysfunctionalities of American cities is their pervasive tendency to split into two distinctive and very unequal sociospatial units. On the one side, we have inner-city areas, endemically given over to poverty, failure, job loss, crime, unemployment, and all the rest; even central business districts are under considerable stress today. On the other side, suburban communities are zones of relative advantage and privilege, with a disproportionately developed capacity to commandeer beneficial public expenditures. But suburbs too impose a heavy toll on urban society due to low-density sprawl and skyrocketing infrastructure costs as the white middle class seeks to flee further and further from the slowly spreading disaster of the inner city. These intertwined problems are exacerbated by the sharp politicoadministrative split that commonly exists between the inner city and the suburbs in the American metropolis.

Metropolitics is an account of Myron Orfield's personal odyssey of political contestation of this syndrome in Minneapolis-St. Paul over the early 1990s. The book opens with an extended description of the usual manifestations of urban decay and decline, all of them heightened in the case of Minneapolis-St. Paul by a particularly elevated rate of inner-city job loss over the 1980s. Perhaps the most exasperating of all of the problems that Orfield enumerates is the desperately poor quality of inner-city schools, for this single deficiency then translates into a multiplicity of other life-long breakdowns. It is all the more exasperating because, as Orfield continually points out, much of it can be traced to the inadequate tax base of inner-city school systems, which in turn is a direct result of the political self-isolation of suburban municipalities. Inner cities, in short, are marked by deeply rooted negative synergies of disadvantage.

Orfield's solution to this dilemma is a thoroughgoing regionalism, meaning political and administrative integration of all parts of the wider metropolitan area. His suggested program of reform involves three core objectives, namely, fair housing, property tax sharing, and reinvestment; and the immediate tools by which he proposes to attain these objectives are systemwide land planning and growth management, welfare reform, and integrated infrastructure financing and management.

How, we may very well ask, is an agenda like this to be achieved in view of the persistent resistance of wealthier white suburbs to any program of intraurban redistribution? Orfield answers this question via a detailed account of his own practical experiences in building a political reform movement in Minneapolis-St. Paul over the period from 1993 to 1996. The key to success, in the Twin Cities and elsewhere, he claims, lies in forging a strategic alliance between inner-city residents and the denizens of older blue-collar suburbs, the latter themselves threatened by imminent decline and neglect. Even so, these two social groups are not usually spontaneous allies, and meticulous coalition building is necessary to induce them to work cooperatively in pursuit of overall change. Orfield documents his own and others' strenuous efforts in the Minnesota House of Representatives and in the Twin Cities to build just such a coalition. One immediate result of these efforts was the reorganization of the Twin Cities' regional planning council from a \$40 million-a-year agency to a \$600 million-a-year governance structure with coordinated control over land use, housing policy, transit and transportation, airports, and sewers across the entire metropolitan area. In addition, Orfield and his political associates secured passage of a regional affordable housing bill and approval of a new regional revenue-sharing formula. All of this legislative activity was accompanied by intensive political mobilization around regional reform issues and was consolidated in 1994 by the emergence of the Alliance for Metropolitan Stability out of a coalition of community and good government groups.

Orfield's arguments are highly persuasive. Without some decisive new political initiatives, the vicious circle of the metropolitan crisis as we

know it is likely to endure indefinitely: Suburban communities will continue to appropriate the benefits of their association with the metropolis while evading the concomitant responsibilities, and inner cities will continue to decay by reason of their own negative synergies. Moreover, attempted market solutions to the problem predictably and persistently fail, as the Rebuild Los Angeles organization found to its cost. Enterprise zones have similarly failed to bear fruit. And the problems deepen further as cities find themselves caught between the pincer movement of the intensifying competition unleashed by globalization on the one hand and the retreat of the federal government from involvement in urban affairs on the other. In this context, as Orfield suggests, the potentialities of state legislatures for constructing a decisive counterattack on metropolitan problems have hitherto been remarkably underestimated.

The fundamental irrationalities of American cities impose enormous costs on society. The wanton squandering of the human resources of the inner cities is an intolerable scandal in the face of so much wealth in the country as a whole; and the concomitant waste of one of the most valuable assets of the metropolis, namely, the centrality and accessibility of inner-city locations, deeply undermines overall regional economic performance. On the grounds of both simple humanity and hard-headed economic calculation, then, major healing of this blighted organism is long overdue. But any such healing can only begin in earnest once the necessary tasks of political activation have been accomplished. This in turn presupposes a campaign of political education able forcefully to expound the theme that current patterns of urban social polarization subvert the ability of the city to reap all of its latent competitive advantages. As I have argued in detail elsewhere ("Regional Motors of the Global Economy," *Futures* 28 [1996]: 391-411) this theme is all the more pressing in the context of the new global economy where regions find themselves competing against one another head-to-head and where creative local institution building is an essential prelude to high levels of regional economic performance and competitiveness.

Orfield's book is one of a number of hopeful signs that this theme is increasingly entering into the political consciousness of regional collectivities. The only problem with the book, I would add, is that it still does not come anywhere near to articulating the radical reconstruction—in principle and in practical politics—of the governance structures of regions that the new world order now makes imperative.

Gangsters: Fifty Years of Madness, Drugs, and Death on the Streets of America. By Lewis Yablonsky. New York: New York University Press, 1997. Pp. xx+237. \$26.95.

Daniel Monti
Boston University

Lewis Yablonsky has been writing about youth gangs and working with gang members since the 1950s. His most recent work on these subjects, *Gangsters: Fifty Years of Madness, Drugs, and Death on the Streets of America*, is best read as a retrospective look at the development of his thinking on gangs over the last four decades. In writing this book Yablonsky joins two other notable figures of his generation—Malcolm Klein (*The American Street Gang* [Oxford University Press, 1995]) and Irving Spergel (*The Youth Gang Problem: A Community Approach* [Oxford University Press, 1995])—who also authored recent books that laid out their views about the changing character of gangs during the same period. As creators and advocates of much of what passes for conventional wisdom on matters related to youth gangs, these three men deserve to be read and taken seriously.

Common to all three works is the idea that communities where gangs appear and the gang members themselves are more deeply troubled today. They are more violent. More of their violence is lethal. They are more deeply involved in criminal behavior, particularly in the distribution and sale of illegal drugs. The young persons who participate in gangs are more likely to start their careers as gangbangers earlier and stay with gangs longer.

Common also to all three books are explanations of gang organization and behavior that focus on the alleged disorganization of the communities where gangs appear and the personal "disorganization" evident in gang members' social and emotional lives. Social disorganization is associated with dramatic changes in the size and composition of certain populations in a community. These populations usually are composed of individuals who are poor or who are minorities of one sort or another. Something in the way of life or values they practice lends many of them to participate in unconventional activities.

Yablonsky, like most gang observers, has his own scheme for talking about different types of gangs. Though no single gang may exhibit all the features deemed applicable to any particular type of gang, Yablonsky (pp. 40–43) believes that gangs in the past generally were "social," "delinquent," or "violent." What is different about contemporary gangs, he maintains, is that they have become "multipurpose." That is, they combine elements of three types of gangs that he says were once pretty distinctive.

The emergence of multipurpose gangs and the kinds of "sociopathic" behavior exhibited by gang members today is unprecedented. It is for this reason, perhaps, that Yablonsky dwells so much on the extreme be-

havioral and emotional problems allegedly exhibited by gang members. He tries to convey his sense of how dangerous and troubling these young persons are, and he generally succeeds. After all, fixing the label "sociopath" to not just one or two individuals but potentially to many thousands of young persons all across the country is no small thing. Not for a moment am I suggesting that Yablonsky uses the term in a gratuitous manner. Rather, it allows him to make his case for subjecting gang members to intensive therapeutic regimens that might save them from committing heinous acts in the future.

Yablonsky talks a great deal about the therapist who works with gang members in small groups. He believes that such strategies hold the best hope for reaching gang members and making them less self-centered and more empathetic and moral. His counterparts, Klein and Spengel, prefer bigger programmatic initiatives and amendatory efforts parachuted into the community where gangs are found.

Yablonsky is not immune to the appeal of these larger social action programs. He argues that a "society that fails to find remedies for its own institutionalized social inequities . . . is likely to continue to suffer with the existence of gang violence" (p. 125). His own work, however, draws him to more modest actions. His "adult youth association approach" builds on older ideas and techniques associated with settlement houses and "detached gang workers." It is designed to bring young gang members into the orbit of conventional adults who can serve as good role models. Yablonsky's detached gang workers would help to coordinate the activities of "indigenous adults" in the neighborhoods where gangs are found. The object would be to "peel" would-be members away from gangs before they became hard-core gangbangers.

This is a far more workable approach to dealing with gangs. Insofar as its ultimate objective is to wean youngsters from gangs, however, it is not likely to enjoy much success. A more realistic goal would be to limit the number and variety of damaging activities they carry out. We have yet to construct a program that will break up gangs or stop young persons from joining them. Yablonsky's proposed solution, while pointing us in a better direction, will not yield the results he expects. It certainly has not so far.

Continuity and Discontinuity in Criminal Careers. By Paul E. Tracy and Kimberly Kempf-Leonard. New York: Plenum Press, 1996. Pp. xvii + 263. \$44.50.

Elmar G. M. Weitekamp
University of Tübingen

Continuity and Discontinuity in Criminal Careers is a remarkable and important book that analyzes the unique and complete official criminal histories of a population of 27,160 (13,160 males and 14,000 females) from

ages 10 to 26 from the 1958 Philadelphia birth cohort. No other data set of this magnitude exists where the representation of race and socioeconomic status are sufficiently well distributed. The analyses of Tracy and Kempf can answer many important questions that criminologists were unable to shed light on before: for example, the transition from juvenile to adult delinquency and, here especially, the continuities and discontinuities of criminal careers for a whole population. This is the first study that can compare the life course development and patterns for a population of males versus females and the crime patterns of adult first-time offenders.

There are several highlights of the careful analyses. Continuity is the main pattern of the 1958 cohort members; most people progressed from a law-abiding life as juveniles to a noncriminal career as adults (71%), while 7.5% of the cohort made the transition from juvenile delinquency to crime. This means that around 78% of the adult cohort members behaved in the same fashion as they had as juveniles. Of the 22% of the cohort members who showed discontinuing behavior, 16% desisted from crime while 1,576 first-time offenders (5.8%) started criminal behavior as adults. With regard to the adult first-time offenders, they found that this group had different criminal careers than the offenders who started their criminal careers as juveniles, and the adult first-time offenders were not a disproportionately frequent or serious group of criminals. With regard to the female-male comparison, the authors found that the ratio for females was much lower. However, differences were found with regard to the offense types, crime specialization, the commitment of violent crimes, adult offending, and co-offending. They found for females no early severity patterns, while such a pattern existed for males.

However, when reading *Continuity and Discontinuity in Criminal Careers*, one always has the impression that underlying the analysis in the book is the concept of the chronic offender. In other words it asks the questions, who are those bad guys or "superpredators" who are responsible for causing so much harm, why do some subjects have long and serious criminal careers, and how can we stop them. It must be pointed out that the definition of a chronic offender, a person who accumulates more than five arrests before his eighteenth birthday, is absolutely arbitrarily chosen and does not make any meaningful distinction. Therefore, policy implications based upon this concept are absolutely meaningless. It is time to write a book about why so many people never become criminals and why so many criminals (even chronic offenders) terminate their careers, and not why only very few persons continue serious criminal behavior. For example, if one looks at the number of "superpredators" in the 1958 birth cohort (defined as those who accumulated five arrests as juveniles as well as adults), we find a total of 229 males: 49 white, 173 African-American, and 7 Hispanic males. This means that 1.7% of the male cohort can be considered "superpredators." It becomes even more ridiculous if one looks at females in the 1958 birth cohort: one white and six African-American females fall into this category, or 0.05% of the fe-

male cohort. This means that in the 1958 birth cohort we find 235 "super-predators" among the 27,160 cohort members, 0.86% of the cohort. It might be worthwhile as well to analyze the temporal order of the offenses rather than the offenders. In "Multiple and Habitual Offending among Young Males: Criminology and Criminal Policy Lessons from a Re-Analysis of the Philadelphia Birth Cohort Studies" (*International Annals* [1996] 34 1/2: 9-53) Weitekamp et al. found in the 1958 cohort that up to the age of 18, more than 50% of the homicides, rapes, robberies, assaults, and burglaries were committed as first, second, third, or fourth offenses and not later in the careers of the cohort members. Paul Tracy and Kimberly Kempf-Leonard clearly have the opportunity to write such a book, especially if one considers that their announced analyses of the follow-up interviews contain an enormous amount of important information that is not available in the official data set.

The policy implications they derive from their analyses are puzzling. One wonders how someone can recommend interventions as early as possible and the use of probation early on if one has no information about what happened after the imposition of probation. Was probation revoked and is this fact responsible for having an effect on the later criminal career? As long as one has no information or can control the intervention effects, it seems to be futile and misleading to base policy recommendations on such analyses. However, I agree, in general, with the policy recommendations of Tracy and Kempf-Leonard that offenders should be held accountable for their actions but recommend that the theories and programs of restorative justice which focus exactly on many things the authors emphasize, be examined.

The Self in the Family: A Classification of Personality, Criminality, and Psychopathology. By Luciano L'Abate with the collaboration of Margaret S. Baggett. New York: John Wiley and Sons, 1997. Pp. xii+404. \$45.00.

Kenneth C. Land
Duke University

The purpose of *The Self in the Family* is to elaborate, update, and expand on a developmental theory of interpersonal competence and socialization previously stated by the author. L'Abate's theory of the self in the family stresses that the family is the major setting in which personality development and socialization take place. The theory proposes that the ability to love and the ability to negotiate—by which *personality* is defined—are the cornerstones of personal and interpersonal competence. *Selfhood* is defined by how one asserts, expresses, and defines his or her importance in intimate as well as nonintimate relationships. This book elaborates on both definitions. Two models of importance and intimacy are expanded in their multifarious functional and dysfunctional manifestations.

The theory then is applied to develop guidelines for psychological interventions relating to prevention and crisis intervention and psychotherapies.

The key word in the title to describe the content of this book is "classification." L'Abate defines a basic diamond-shaped classification of selfhood, which then is repeated over and over again throughout the book as he applies the classificatory theory to various aspects of personality, psychopathology, and criminality. At the top vertex of the basic selfhood diamond is "selffulness" and at the bottom vertex "no-self"; to the right vertex of the diamond is "selflessness" and to the left vertex is "selfishness." Applied to selected personality patterns, he then places "interdependence" near selffulness at the top of the diamond, "overdependency" near selflessness to the right, "extremes in dependency and in denial of dependency" near no-self at the bottom, and "denial of dependency" near selfishness to the left. Applied to psychopathology and criminality, L'Abate places "functionality" at the top of the diamond, "depressions" to the right, "psychopathology" at the bottom, and "criminality" to the left. Within the diamond, he then locates "adequate personality" above the horizontal line connecting left and right vertices and "adjustment disorders" below this line. Various disorders from the *Diagnostic and Statistical Manual IV* [American Psychiatric Association, 1994] and crime categories then are located around the diamond—with "creative personalities" near the top vertex, "anxieties, phobias, and depressions" near the right, "murder and suicide" near the bottom, and "nonviolent and violent crime" near the left.

The book contains two chapters reporting empirical evidence "corroborating the theory." The first surveys a selection of studies in psychology that provide "independent and indirect evidence" of the empirical veracity and applicability of the concepts, classifications, and propositions around which L'Abate's theory is developed. For instance, a recent study of children and adolescents reviewed in this chapter found four major dimensions that make up a "taxonomy" of behavior styles: (1) adjusted and adequately adjusted (characterized, e.g., by inhibition), (2) marginally adjusted (characterized, e.g., by withdrawal), (3) at risk (characterized, e.g., by undersocialized aggression), and (4) maladjusted (characterized, e.g., by instrumental aggression). L'Abate suggests that these dimensions bear similarities to the selfhood model: selffulness for (1), selflessness for (2), selfishness for (3), and no-self for (4). The second corroborative chapter on "direct evidence" similarly reviews a number of studies of the selfhood model conducted by L'Abate and various collaborators. For example, one such study evaluated the relationship between the selfhood model of personality propensities and depression: The model predicts that individuals who rate themselves high on both self and others (i.e., selffulness) would have the lowest levels of depression, those who rate themselves low on both self and others (i.e., no-self) would demonstrate the highest levels of depression, and those in the other two categories would demonstrate intermediate levels of depression. In a study of 155 under-

graduate volunteers reviewed in this chapter, it is concluded that these predictions are corroborated.

The present reviewer claims no expertise in psychiatry or clinical psychology and thus cannot evaluate much of the content of this book. Two characteristics of the volume are striking, however. First, the book provides evidence that the state of personality/self theory has not progressed to a sufficient stage of accumulation such that the theory developed by L'Abate clearly builds upon, and extends, an accepted model. Rather, L'Abate reviews other major theories of personality, lists their shortcomings, and proceeds to develop his own theory. Second, while the theory developed by L'Abate cites many extant theories and empirical studies, the boundaries of the literature cited extend only through psychology and psychiatry. The body of theory and studies of self done within the discipline of sociology or from a sociological social psychological perspective over the past several decades simply are not noted or integrated into the volume. This is not to say that the theory developed by L'Abate is irrelevant to sociology or sociologists. Certain propositions, such as those relating various aspects of personality to criminality, should well be of interest to, and subjected to test by, sociologists who study crime and deviance. It is unfortunate, however, that contributions by sociologists to the theory of the self are not acknowledged and integrated into the present volume.

Homicide, Race, and Justice in the American West, 1880-1920. By Clare V. McKanna, Jr. Tucson: University of Arizona Press, 1997. Pp. xiv+206. \$40.00.

Richard Block
Loyola University of Chicago

Homicide, Race, and Justice in the American West is an empirical study of lethal violence in three Western counties. It combines coroner's inquest reports, indictments, trial and appeals court records, and news accounts to simultaneously analyze the statistics and the nature of homicide and racial/ethnic discrimination in the legal system in the American West. While historical studies of lethal violence using similar methodologies are not new (see Ted R. Gurr, Peter N. Grabosky, and Richard C. Hula, *The Politics of Crime and Conflict: A Comparative History of Four Cities* [Sage, 1997]; or Dirk Arend Barents, *Misdaad in de Middeleeuwen* [Criminality in the Middle Ages] [Utrecht, 1976]). McKanna argues that historical studies of violence in the American West have been mainly anecdotal. This book also has many anecdotes and stories, but it also contains statistical analysis of each county.

While the book is a good read and informative, it is flawed in several ways. First, the choice of counties is peculiar and not very representative. Douglas County (Omaha), Nebraska, grew very rapidly in the late 19th

century with a significant influx of migrants from Europe and the South to man the meat-packing industry. It grew quickly during this time period but in many ways was a part of the Midwest corn belt rather than the West. Las Animas County (Trinidad), Colorado, was an isolated mining area that grew rapidly in the early 20th century with an influx of miners from Europe. It was the site of the Ludlow massacre in 1914, one of the bloodiest events in American labor history. Gila County (Globe), Arizona, was another mining area that also had an influx of European immigrants. Its main distinguishing characteristic was the enmity between European and Apache residents of the county. All three counties are interesting, different, and somewhat unique, but these elements, which make for good reading, also make comparison or the development of a model of lethal violence impossible.

McKanna argues that racial and ethnic demographics are important in explaining the high levels of homicide in each of these counties. Even though census data were available, he almost entirely ignores age and gender demographics. Older residents are barely mentioned, and women are primarily associated with saloons and bordellos. The age and gender structure of frontier communities was heavily weighted toward young men with a relative absence of women, children, and the elderly. Age and gender distributions probably had as much to do with high levels of violence in the American West as racial and ethnic variation.

The author argues that a subculture or culture of violence permeated these counties and led to high levels of lethal violence. He contends that a subculture of violence among Omaha blacks and a culture of violence in Las Animas and Gila counties resulted in high levels of violence. Blacks in Omaha had recently emigrated from the South and carried with them a fear and distrust of white authority along with a Southern tradition that took personal insults very seriously. These factors combined with high levels of alcohol consumption and the ready availability of handguns and led to high rates of homicide (see Raymond D. Gastil, "Homicide and the Regional Culture of Violence," *American Sociological Review* 36 [1971]: 412-27). To what extent the Southern culture of violence is attitudinal or structural has been widely disputed (Colin Loftin and Robert H. Hill, "Regional Subculture and Homicide: An Examination of the Gastil-Hackney Thesis," *American Sociological Review* 39 [1974]: 714-24). This dispute, while recognized in a footnote, is ignored in the text.

In the two mining counties, rates of homicide were high for every ethnic group, thus there is a culture rather than a subculture of violence. Alcohol and concealed handguns were also part of these cultures but so were labor/management conflict, company towns, immigrant customs in Las Animas County, and hatred and distrust of Apaches in Gila County. The explanation of high levels of violence in each county are unique but ultimately based on culture. However, the book flounders in its discussion of the meaning and development of culture in the American West.

Screening the Los Angeles "Riots": Race, Seeing, and Resistance. By Darnell M. Hunt. New York: Cambridge University Press, 1997. Pp. xv+313. \$59.95 (cloth); \$19.95 (paper).

Richard M. Merelman
University of Wisconsin—Madison

Darnell M. Hunt intends *Screening the Los Angeles "Riots"* to bridge a gap between critical media studies and the sociology of race. To this end he deploys both qualitative and quantitative analyses, micro and macro theoretical perspectives, and a formidable—sometimes excessive—array of recent social theory. The target of this analysis is the way differently "raced" groups—Latinos, blacks, and whites—interpreted media coverage of the 1992 Los Angeles "riots."

Hunt's primary concern is the power struggle between audiences and broadcasters over the meaning of disruptive public events. Dismissing the claim of media "objectivity," Hunt argues that television disseminates hegemonic assumptions that support dominant social groups. The two significant theoretical questions he explores are whether hegemony is seamless (as in Althusser) or unstable (as in Gramsci) and whether audiences accept or oppose the hegemonic assumptions embedded in television discourse.

To explore these questions Hunt examines the way small groups of friends respond to a 17-minute excerpt of local television news coverage about the disorders following the Rodney King verdict. Hunt first identifies 14 hegemonic "core assumptions" in the televised excerpt. Next Hunt examines the way in which 15 small friendship groups—five white, five Latino, and five black—respond to this excerpt. He combines linguistic analysis, ethnographic observations, systematic comparison of group conversations about key events on the tape, and attitude scales administered to viewers. Hunt examines both the content of group conversations and modes of group viewing. He reasons that passive "referential" viewing and expressions of viewer agreement with the core assumptions of the tape indicate acceptance of the tape's hegemonic frame. Animated "metalinguistic" viewing and opposition to the tape's core assumptions indicate potential resistance to hegemony.

Not surprisingly, Hunt finds that it is the black groups that manifest most opposition to the tape's core assumptions. Latinos weakly and whites strongly support the hegemonic assumptions of the tape. A full 73% of all groups decoded the tape referentially rather than metalinguistically. Two of the 14 core assumptions of the tape received sustained opposition by all three raced groups, though these two assumptions—the undesirable nature of the events and the fairness of media coverage—are quite important. Finally, on one of three attitude scales—the "arrest" scale—there was significant racial group polarization as a consequence of viewing. Hunt concludes that "study groups generally made negotiated decodings of the . . . text. That is, they decoded the text according to

various mixtures of dominant and oppositional codes" (p. 161). Hunt believes this negotiated meaning contains the seeds of resistance to media hegemony.

Hunt should be commended for subjecting aspects of critical media studies to empirical investigation. Moreover, in chapter 7, he breaks new ground by identifying some of the linguistic methods people use to make sense of media content. These are real contributions.

However, Hunt's conclusions are debatable. Relying upon a few small groups of college-age people clearly limits the scope of conclusions. Because these groups gathered together primarily for the viewing occasion and were then prodded to discuss what they were seeing, the setting is somewhat artificial. Moreover, several of the "raced" groups are in fact racially mixed. Finally, we have no way of knowing whether the taped excerpt is representative of Rodney King news coverage.

More important, Hunt's findings do not support his claims about opposition to the text's hegemonic assumptions. Both before and after viewing and while discussing the tape, Latinos and whites were overwhelmingly opposed to the looting and arson depicted and overwhelmingly favorable toward the arrest of event participants. Blacks were ambivalent and divided on these same issues, though they did become more opposed to the arrest of event participants. White and Latino groups also accepted 12 of the 14 core assumptions of the media excerpt, while blacks questioned only four. Finally, only when specifically prompted to do so did most study groups question the assumption of objectivity in the excerpt. On balance, the combination of white/Latino espousal of hegemonic positions and black ambivalence toward those same positions seems a recipe for continued group domination, not group resistance.

Although Hunt dismisses media claims of objectivity out of hand, the television excerpt actually contains some counterhegemonic as well as hegemonic content. For example, one newsman describes the events as an expression of community frustration, rather than as a "riot." In another place a black interviewee argues that the police *want* blacks to destroy their own neighborhoods and then urges blacks to defeat this police tactic. Perhaps the negotiation of meaning between media and audience is more complicated than the account that Hunt presents. Hunt's book is a halting first step in an important direction; let us hope more successful studies follow in its wake.

White Lies: Race, Class, Gender, and Sexuality in White Supremacist Discourse. By Jessie Daniels. New York: Routledge, 1997. Pp. xvi+171. \$55.00 (cloth); \$15.95 (paper).

Noel A. Cazenave
University of Connecticut

White Lies is part of an important intellectual movement to place white racism at the center of social science and literary analyses. That movement is fueled by the success of both African-American and lesbian feminists in forcing the feminist movement to integrate treatments of race, class, and sexuality into gender studies and by the growing body of literature on the social construction of whiteness.

The book's most compelling argument is that studies of white supremacist discourse are important, not simply for what they reveal about white supremacist extremists but for what they can tell us about the ideological glue that holds together this nation's social center. A promise of such analyses is the light they can shed on the quiet everyday workings of ordinary institutionalized white racism. Daniels argues that the social construction of whiteness and the race, class, gender, and sexuality inequality functions served by white supremacist discourse are central to such disquisitions.

Daniels' study is based on ethnographic content analysis of 369 issues of six white supremacist publications for various blocks of years ranging from 1976 through 1992. *White Lies* is organized into six chapters covering past and present white supremacist movements, their social functions, and their treatment in the social science literature; white supremacist depictions of white, black, and Jewish men and women; and an exploration of various strategies to combat white supremacist groups.

White Lies is appropriate for courses in race, class, gender, and gay and lesbian studies as well as American studies. I plan to use it in both my undergraduate and graduate classes on white racism. This book should also prove useful in courses on the social construction (or phenomenology) of whiteness.

Daniels' tendency not to expand on the provocative assertions that she makes is evident in her treatment of Jews in white supremacist discourse. She states that, to the extent that Jews in the United States have successfully adopted a white racial social identity, they have gained protection against anti-Semitism. This assertion is significant because it implies that American Jews would find it especially difficult to relinquish white racial identity. The validity of this argument and its implications for Jewish social identity, ethnic group relations, and the growing movement in the United States to question, or indeed, relinquish white racial identity merit further exposition.

While Daniels' book effectively debunks numerous "white lies" prevalent in the United States today, it fails to challenge the ultimate lie of

"whiteness" itself. Consequently, *White Lies* fails to adequately meet one of its core objectives—to increase our understanding of the social construction of whiteness. The achievement of that goal requires more historical analysis than Daniels offers and a theory of white racial hegemony and its key ideological components—"race" and "whiteness." In the absence of such an approach, Daniels is unable to explain either the origins of white supremacist discourse or how it is structured within contemporary American mainstream society. Consequently, Daniels is not equipped to achieve her goal of developing a method capable of "dismantling rather than sustaining, systems of domination" (p. 144).

The inability of the book's analysis to obtain its author's radical goal is clear in the few pages Daniels devotes to strategies of combating white supremacy. Daniels cavalierly dismisses the New Abolition Movement's call for the development of an antiracist movement predicated on the abolition of white racial identity. Daniels's skepticism seems to be based on no more than her concern that such change could be unsettling. As she puts it, "it is not clear what it means to *not be white*, or even to begin to move in that direction" (p. 137). Indeed, Daniels seems unconvinced (by her own study) that white racial hegemony is a central and enduring principle around which the United States and other modern societies are structured and evolve. This is the essence of Derek Bell's "permanence of racism" argument, which she all too quickly dismisses as "still an open question" (p. 137).

In the end, because of Daniels's failure to come to grips with the fact that effective antiracist analysis and activity must go beyond "dealing with race" (p. 137), *White Lies* fails in its mission to pose a fundamental challenge to white racial supremacy. As long as America's core ideological constructs of "race" and "whiteness" remain in place, Daniels's goal of dismantling its system of white racial hegemony will be unattainable. In brief, *White Lies* is a good book, as far as it goes. Unfortunately—in its exploration of the organization and persistence of white racial supremacy—it does not go far enough.

Teen Mothers and the Revolving Welfare Door. By Kathleen Mullan Harris. Philadelphia: Temple University Press, 1997. Pp. xv+195. \$39.95.

Sandra K. Danziger
University of Michigan

In *Teen Mothers and the Revolving Welfare Door*, Harris insightfully analyzes how a major group within the "welfare population" uses public income support and how welfare reliance interacts with other aspects of women's lives, such as family formation, schooling, and work. She creatively and carefully replicates policy analyses that have influenced current political debates over welfare reform. She shows that even though

urban African-American teen mothers in this study conform to certain underlying assumptions about welfare dependency, the differences among them challenge most prevailing stereotypes.

Most teen mothers in this study receive welfare benefits at some point during the 20 years in which they were followed. However, they also accumulate substantial work experience, and those who formulate an "education strategy" at an early age fare best economically throughout their later lives. Subsequent economic outcomes have little to do with socialization into a welfare career or how quickly they marry or start to work. In fact, Harris finds, surprisingly, that early marriage and early labor market experience may negatively affect long-run potential to stay off the welfare rolls.

The book clearly explains the event-history techniques utilized and their importance for addressing policy questions and real-world issues. Harris builds the rationale for this model in each chapter and also provides case illustrations to illuminate the pathways that result in welfare entry, exit, and return. Her unique sample, the Baltimore teen mothers who gave birth in the 1960s and were interviewed several times by Frank Furstenberg and colleagues, provides 20 years of rich life-history information.

Harris contextualizes her study within the literature on poverty and the urban underclass, the history of welfare policy approaches and their sociological research analogs, and shows how studies of the consequences of adolescent parenthood fit within these traditions. She dispassionately reviews these debates and their theoretical and empirical underpinnings. She frames her study narrowly within the research on outcomes of teen pregnancy and the question of whether poverty and poor economic life chances result from the same conditions that increase the risk of early birth or from the teen birth itself. This distinction, however, cannot be resolved with her sample because it does not include women who defer childbearing until later years. Instead, she asks which of the teen mothers overcome their family background and age-at-birth characteristics and who among them fail to improve their economic status as the children grow up. Specifically, how do these early factors condition later welfare-related experiences in comparison to continuing life events and cumulative experiences?

The analyses rely on the life-history calendar recordings that subjects filled out and that represent 20 annual data points for welfare use, employment status, and marital, childbearing, and schooling events. Although Harris explains how the measures were constructed and analyzed, a few limitations could have been better addressed. First, she uses year-to-year changes to categorize spell beginnings or endings. She reports that a large number of welfare changes were not associated in time with employment, schooling, marital, or educational events. For example, 20% of the returns to welfare had "unknown" types of spell beginnings. Harris makes a plausible case for one interpretation of these welfare shifts—that

they result from informal income changes not picked up in the calendar. However, the annual life history calendar itself may underreport some of the very changes she examines, particularly employment changes. Many entry-level minimum wage jobs turn over quickly, and a measure of whether a respondent worked at any time during a year could miss job changes or disruptions.

Similarly, her analysis of women who combine work and welfare is limited by this measurement problem. Harris finds that the longer a person remains on welfare, the greater her probability of working at some point in the same year as she received welfare. This is similar to finding that getting beyond one's teen years or having one's youngest child age increases the chances of entering the labor force. She documents that working is more common for these women than is its total absence, but the indicator does not reflect the type or degree of work activity.

Her major contributions are that these women appear quite similar to other poor single mothers and that, although they entered adulthood decades ago, in a different policy era and under different poverty conditions, their paths into and out of welfare dependency have important policy significance today. She raises the specter that just leaving welfare within a certain time frame does not predict well-being at a later point, that those who cycle on and off the rolls may be worse off than those who stay for a longer period of time but who avoid subsequent welfare spells.

This book should be read and debated in family sociology and family policy classes, in policy analysis and research methods classes, and it deserves a full hearing by legislators and welfare administrators. Gratitude goes to Frank Furstenberg and his colleagues for generating these longitudinal data treasures and to Kathleen Harris for this timely, thoughtful, and well-crafted study.

Immigration and the Family: Research and Policy on U.S. Immigrants.
Edited by Alan Booth, Ann C. Crouter, and Nancy Landale. Mahwah,
N.J.: Lawrence Erlbaum Associates, 1997. Pp. viii+307. \$59.95.

Nancy Foner
State University of New York, Purchase

The family, Rubèn Rumbaut observes, is a strategic site for understanding the dynamics of immigration flows and immigrant adaptation, yet it is an area that has received relatively little attention in studies of recent immigrants in the United States. This volume seeks to redress the balance. It is based on a 1995 conference on international migration and family change at Pennsylvania State University that brought together a group of sociologists, psychologists, and policy experts to examine four main questions: who migrates and the effects on family outcomes, the impact of the migration experience on child and adolescent development,

the way family processes and structure change across succeeding generations, and whether government policies enhance or impede immigrant family links to U.S. institutions. The book follows a conference format: each of the four sections includes a lengthy lead-off paper on one of the main questions, followed by two or three briefer commentaries.

Rubén Rumbaut does a good job setting the stage by providing a profile of the new immigration in terms of class and national origins. He discusses the role of family networks and ties in understanding who migrates—over three-quarters of immigrant visas now go to family members—and determining where the new arrivals settle. Drawing on recent survey research on the children of immigrants in California and Florida, he suggests some ways that family contexts interact with human capital to influence children's educational achievements. His emphasis on the role of "strong family solidarity centered on a cohesive conjugal unit and parental bond" in immigrant children's success, especially in the face of limited human capital and discriminatory conditions (p. 28), points to the need for further studies to identify the dynamics involved.

Mary Waters's commentary paper is an excellent example of the way ethnographic work on immigrant families, in this case West Indians, can shed light on the multiple factors influencing children's educational outcomes. She shows how a complex combination of social factors puts strains on immigrant families and may contribute to a decline in aspirations, academic achievement, and hard work among immigrant children. These include serial migration, the isolation of nuclear families, the long hours parents spend at work, conflicts between the immigrants and American norms about discipline for children, racial discrimination, and the disinvestment of American society in inner-city neighborhoods and schools.

Two lead papers offer reviews of the literature. One, by Cynthia Garcia Coll and Katherine Magnuson on the psychological effects of migration on children and adolescents, discusses the sources of, and responses to, stress and conflict, though the need to consider the possible benefits of migration is noted as well. Raymond Buriel and Terri De Ment describe changes in family structure, childrearing, intergenerational conflict, and family roles in light of empirical evidence from the research literature on Mexican-, Chinese-, and Vietnamese-American families. They argue that a bicultural orientation emerges among many immigrants, with acculturation to mainstream American culture coexisting with the retention of aspects of the homeland culture. As Nazli Kibria points out, questions remain about the nature of the intermingling between cultures and how the different cultural strands actually operate in relation to each other. And, as both Kibria and Charles Hirschman indicate, immigrants' homeland cultures—as well as mainstream American culture—are more diverse than Buriel and De Ment imply.

United States government policies and laws can have far-reaching effects on immigrant family relations, often in unintended ways. In consid-

ering the impact of proposed immigration policies and welfare reforms, Michael Fix and Wendy Zimmerman note that if restrictions are placed on the admission of extended family members, the social capital of immigrants will be reduced. In addition, by dramatically curtailing most non-citizens' access to means-tested public benefits, recent welfare reforms will increase financial difficulties facing immigrant families and create incentives to naturalize as a way to retain eligibility for benefits.

Given the broad nature of the topic, it is inevitable that some themes related to immigration and the family are left out or given scant attention, one example being changing gender relations. In the concluding chapter, Nancy Landale (one of the editors) notes that the volume only briefly addresses questions related to the way the family itself changes as a result of immigration. Clearly, there is a need for additional studies that examine the way immigrant families change in response to structural, cultural, and economic forces in this country—and on the consequences of new family forms, relations, and values for immigrants and their children. Happily, the volume provides a useful review and assessment of many research findings in the field and raises important questions for further study. Those concerned with the role of the family in the experience of immigrants in the United States will definitely want to read it.

The Modernization of Fatherhood: A Social and Political History. By Ralph LaRossa. Chicago: University of Chicago Press, 1997. Pp. xi+287. \$18.95.

E. Anthony Rotundo
Phillips Academy

Ralph LaRossa has thrust himself energetically into the search for true historical fatherhood which has bedeviled recent scholarly and political debates. He maintains that current advocates and exemplars of "New Fatherhood" understate the degree of father involvement in the past in order to highlight their own achievements and concerns today. LaRossa also asserts that scholars creating the emerging history of fatherhood have denigrated fatherhoods past by ignoring some historical data and making rash assumptions about others.

The heart of LaRossa's argument is that historians, who have typically placed the emergence of modern fatherhood at the turn of this century or in the past few decades, are wrong. Our current concept, he says, was a product of the 1920s and 1930s. To support this novel assertion, LaRossa explores institutions, movements, and individuals that both created and reflected the modern culture of fatherhood, notably the Children's Bureau, the child-study movement, *Parents' Magazine*, the parenting columnist Angelo Patri, and the movement to celebrate Fathers' Day.

In truth, I have given LaRossa's argument a clearer statement than he

has. The book's chief weaknesses are the rhetoric, focus, and definition. LaRossa variously describes his subject as "the current image of the father," "the modernization of fatherhood," and the "New Fatherhood." He tends to use these terms interchangeably, and, yet, he defines them differently. The current image of the father consists of "economic provider, pal, and male role model all rolled into one" (p. 3), while the modernization of fatherhood (which, in spite of being the book's title, is not defined until pp. 68-69) "was basically about making men more caring and more involved." Thus, the current image of the father and the modernization of fatherhood are distinct sociocultural phenomena, however closely related. Actually, LaRossa uses the New Fatherhood designation, a label he never defines at all, more than either of the other two terms.

LaRossa fails to lead the reader through the subtle differences and interrelations among these concepts. Instead, he shifts focus from one to another without warning, while using the terms synonymously. Likewise, LaRossa needed to use a more subtle hand in shunting the reader back and forth between changes in cultural attitudes about fatherhood and actual fathering behavior. In general, his argument with historians is about attitudes, and his dispute with current ideologues of fatherhood is about behavior. But, rather than separate the arguments or narrate two distinct but intimately related histories of idea and practice, LaRossa jumps from one to the other as they pop out at him from the evidence.

Finally, in spite of LaRossa's focus on the 1920s and 1930s, his own evidence points variously to the 1930s, the late 1930s, and the early 1940s as times when patterns of image or behavior changed. And he himself admits that key cultural changes happened before the 1920s: "America's philosophy of fatherhood . . . is firmly planted in soil tilled sixty to seventy years ago *and before*" (p. 10; my emphasis).

Nonetheless, I would advise the reader to wade patiently through the conceptual turmoil because underneath the chaos, this is a book of real value. The greatest value lies in three significant contributions. One is that LaRossa sifts his evidence with a subtlety that all of us involved in the fatherhood debate would do well to emulate. For example, he is admirably persistent in asking what ideological and self-interested filters shaped the views that appear in each of his categories of evidence. He is especially good at distinguishing varieties of gender bias in this manner.

A second important contribution lies in LaRossa's sophisticated analysis of the way cultural ideals are transmitted and then lead (or fail to lead) to changes in behavior. He offers a keen critique of the therapeutic-culture model which holds, among other things, that a child-rearing intelligentsia has determined parenting behavior in the 20th century. LaRossa notes that the advice of the experts is prescriptive but not determinative. People interpret such advice in highly selective ways and often choose to ignore it. LaRossa is also a discerning analyst of the gender politics of fatherhood and of parenting advice. He observes that, in the context

of the still-patriarchal family of the 1920s and 1930s, fathers had great power to reject or revise the parenting prescriptions that their wives might favor. At the same time, he notes that the counsel of child-rearing experts could as often reach fathers through their wives as through direct exposure, and he argues poignantly and persuasively that changes in women's gender-role attitudes constitute a powerful pressure on their husbands to change their own attitudes.

Finally, LaRossa succeeds in making a case for the importance of the 1920s and 1930s in the history of fatherhood. While his own evidence does not support his position that the "modernization of fatherhood" happened between the world wars, he makes a convincing case that previous historians of American fatherhood have erred seriously in simply dismissing the importance of that era.

A History of the Family. Vol. 1, *Distant Words, Ancient Worlds*. Edited by André Burguière, Christiane Klapisch-Zuber, Martine Segalen, and Françoise Zonabend. Cambridge, Mass.: Harvard University Press, 1996. Pp. x+713. \$39.95.

A History of the Family. Vol. 2, *The Impact of Modernity*. Edited by André Burguière, Christiane Klapisch-Zuber, Martine Segalen, and Françoise Zonabend. Cambridge, Mass.: Harvard University Press, 1996. Pp. vii+585. \$39.95.

Rosanna Hertz
Wellesley College

A History of the Family is intended to highlight what the study of family can contribute to our knowledge of a given society and how various historical periods add to a broad understanding of family life. Most of these chapters rely heavily upon the work of social anthropologists—and rarely mention the scholarship of sociologists (which is most markedly absent in the U.S. contemporary chapter). These volumes resemble an exercise in recalling what the "fathers" of social anthropology had written since most of these authors frame their historical material by drawing heavily upon the observations, predictions, and theories of anthropology. There is no attempt to create a historical model of "the family" or "families." Moreover, there is no consistent set of topics linking the chapters to help the reader find such a theory. Instead, the 30 chapters are written with varying levels of detail and depth depending upon the unique focus of each author and, I assume, the available evidence.

Volume 1 focuses on ancient societies, Europe in the Dark Ages and medieval times, and the classical periods of non-European civilizations. Volume 2 examines the impact of modernity with a specific emphasis on the religious, political, economic, and social spheres. Volume 2 also explores how selected colonized regions confronted the dominant European

model of family life. While it is more difficult to assess volume 1 because evidence is more limited, volume 2 appears to present normative or ideal family forms with less emphasis on family diversity or cultural or racial differences despite the fact that more research is available. Instead, social class and rural/urban distinctions between families are more likely to be featured.

I approached these volumes as tools to be used by a sociologist to provide historical material on contemporary topics. For example, as a contemporary qualitative scholar of the family, I accepted this assignment because I believed that these two volumes would provide historical precedents for my present coauthored research on adult women who parent children out of wedlock. At the very least, I hoped to learn something about how different cultures in various historical periods dealt with out-of-wedlock births. Scattered throughout these two volumes are occasional glimpses of that type of information. Basically, they occurred in discussions of inheritance and the distribution of land—for instance, in ancient Athens, illegitimacy was a greater disadvantage to men than to women because the denial of paternity stripped a boy not only of his right to succession but also deprived him of a political life (the “freedom of the city”). Since a girl never shared the same rights as a boy, an illegitimate daughter would suffer only on the level of private law (1:20). Certainly women in all historical periods and countries had children out of wedlock. Sometimes, as in Assyria, the children were claimed by the husband who may have been absent for various reasons (1:115). Yet central questions, which move beyond the legal transfer of property or wealth and the social description of family lineages, such as how the mothers of illegitimate children were treated and how they survived economically, are not answered. More in-depth material on this topic is either completely ignored or inadequately represented in these volumes.

The one exception I could find was a discussion in the chapter on contemporary Scandinavia and the welfare state. Recent changes in household composition and marriage led the authors, David Gaunt and Louise Nyström, to at least raise the question of men’s place in family life. Do women and children really need husbands and fathers? Since the anti-authoritarian parent is the ideal, would new father “requirements,” which make men less disciplinarian and more nurturing and forgiving (like mothers), create havoc and identity crises for their offspring (2:484)?

I suppose that some of the more brutal practices mentioned in these volumes occurred because of out-of-wedlock births, particularly abortion, infanticide, and child exposure. But the question of under what conditions these practices are linked to illegitimate births versus common practices within marriage is not explored. It is possible that historical records that distinguish family form or give sufficient detailed information on the reasons are nonexistent. Instead these practices are treated in the course of discussions about family life, most often as substitutes for contraception. Exposure and incest are practices sometimes legally condemned but socially sanctioned; other times they are quite common occurrences. For

example, in Assyria, abortion was not only condemned but could be punished by impalement and the refusal of a burial place as, the authors speculate, infanticide and child exposure also must have been (1:115). During the Edo period in Japan, abortion and infanticide were common practices (1:553). Under the Roman Empire, brutal practices were part of the fabric of ordinary life. In the family, fathers decided the life or death of children; sickly or deformed children whose fathers did not wish to bring them up could be drowned or exposed (1:295).

The volumes do offer the reader a fresh look at family life. In the introduction to volume 1 Claude Lévi-Strauss writes:

It is no longer possible to believe that the family has taken a unilinear path as it has evolved from ancient forms which we will not see again, moving towards other, distinctive forms that represent progress. On the contrary, it could be that at a very early stage the human mind, in all its inventive power, conceived and laid out on display all the forms assumed by the family. What we regard as a process of evolution would then be nothing more than a series of choices from these possibilities, the result of movements in various directions within the limits of an already defined framework. (1:6)

Françoise Zonabend's article follows as an additional opening to the volume (1:8–68). Featured in this piece is an interesting discussion of the contribution of social anthropology to mapping kinship arrangements from patrilinear models to matrilinear models complete with fascinating examples. Here I could imagine the threat my single mothers would have posed in ancient societies. They would have challenged the existence of one basis of family—that being “exchange relationships”—where families swap members in order to form essential alliances. However, the present rise of single mothers throughout the world cannot simply be accounted for by the lack of exchanging women as part of a community's culture.

The chapters in volume 1 trace the evolution of the family from primate behavior to codified practice in Sumer and Babylon, the ancient civilizations of Egypt, Athens, Rome, and the Hebrews to feudal Europe and medieval Byzantium, concluding with the early histories of China, India, and Arab Islam. These chapters synthesize huge amounts of detail, legal rulings, literature, and demographic materials. This volume left me with a fresh insight on how the vast majority of people did not exist historically as individuals but were deeply embedded in family, kinship, and community. Community was not a separate entity from family but was *of* family and kin. Codified practices of family responsibility and obligations to blood and adoptive kin as well as to the community leave me wondering about the contemporary idea that family is a “private” institution. Privileged classes enjoyed more material wealth and land and often times had servants. Lower classes fared less well. However, all women were defined by their spouses though higher-status women did exercise more authority. For example, in Egypt during the time of the Middle Kingdom on, pharaohs were the primary authority figures, but some of their wives did enjoy considerable authority acting in their ab-

sence and participating alongside them in political circumstances. But the authority of less prominent wives was relegated to domestic and child-rearing tasks. However, some women, the wives of high-status men, were engaged in professional activities and relieved of household tasks, which were most likely done by servants. Interestingly, some sources report that widowhood allowed women to take over as heads of households; other conflicting sources say they returned to their fathers' houses and the children became the responsibility of maternal uncles (see 1:149–52).

Volume 2 looks at the changes in Europe and other countries with the advent of modernity (i.e., the development of writing, the rise of the state, and the expansion of industrial capitalism, colonization, and urbanization). I particularly found the chapters on India and Africa most interesting in illuminating how colonization transformed family life. For instance, colonization created some ethnic entities that had no real equivalents in the precolonization period (2:306). The chapter on India looks specifically at the differential transformation of women and the family in rural and urban areas. The chapter on Africa is a critique of Radcliffe-Brown and, while it points out his errors in understanding and the use of his conclusions for colonization purposes, the chapter could have differentiated more between various regions of this vast continent.

The authors in this volume do not intend to write the family's epitaph, but instead they demonstrate how the family is the cornerstone of all societies in all historical periods, with families remaining embedded in broader structures including kinship and community. Universal matrimonial norms emerge due to the profound influence of the church followed by the privatization of marriage and choice of spouse. Yet, the meaning attached to marriage and its dissolution varies widely. Divorce and remarriage and the reduction in the size of the household and number of children arise around the world despite differences in political systems and religious beliefs. Children today are more likely to have their grandparents than in earlier generations due to increasing life span. Tensions between generations because of proximity and dependence that characterized relationships in rural areas are less tension filled because in contemporary Western societies pensions and social security make the elderly less dependent on the younger generation. Yet, in other cultures such as China and Japan, children still care for their parents. "Working women" is also a theme that most authors address. In rural and preindustrial economies, women did the domestic work and worked in the family enterprise or outside the home in manufacturing or domestic service. In early industrial societies, women worked outside the home in varying jobs including textile manufacturing. Even with the rise of middle-class ideology specifying women as the creators of family life and men as the breadwinners, labor force participation has dramatically increased for all women, not only among women in the working classes but among middle-class women as well. Mostly they are relegated to salaries that are not equal to those of their male peers; and often they work part-time or in lower-

status jobs. Consistent across regions of the world is the "double-duty model": after working a paid job, women are then expected to be responsible for running the house as well. While the editors note in the introduction that these volumes are not meant to be an encyclopedic survey of various phases of family life, the best use of these volumes for a sociologist is as a valuable and well-written sourcebook for selected regions of the world.

CONTENTS

- 1 More Shock than Therapy: Market Transition, Employment, and
Income in Russia, 1991-1995
THEODORE P. GERBER AND MICHAEL HOUT
- 51 A Frame Analysis of Favor Seeking in the Renaissance: Agency,
Networks, and Political Culture
PAUL D. MCLEAN
- 92 Toward a Role-Theoretic Conception of Embeddedness
JAMES D. MONTGOMERY
- 126 Rethinking Age Dependence in Organizational Mortality: Logical
Formalizations
MICHAEL T. HANNAN
- 165 Uncertainty, Trust, and Commitment Formation in the United
States and Japan
TOSHIO YAMAGISHI, KAREN S. COOK, AND MOTOKI WATABE
- 195 Status Processes and Mental Ability Test Scores
MICHAEL J. LOVAGLIA, JEFFREY W. LUCAS, JEFFREY A. HOUSER,
SHANE R. THYE, AND BARRY MARKOVSKY
-
- Book Reviews**
- 229 *Coffee and Power: Revolution and the Rise of Democracy in Cen-
tral America* by Jeffery M. Paige
WILLIAM ROSEBERRY
- 231 *Restructuring Networks in Post-Socialism: Legacies, Linkages, and
Localities*, edited by Gernot Grabher and David Stark
ÁKOS RÓNA-TAS
- 233 *Contemporary Capitalism: The Embeddedness of Institutions*, ed-
ited by J. Rogers Hollingsworth and Robert Boyer
ALBERTO MARTINELLI

- 235 *Eco-Nationalism: Anti-Nuclear Activism and National Identity in Russia, Lithuania, and Ukraine* by Jane I. Dawson
DAVID R. MARPLES
- 237 *Russian Talk: Culture and Conversation during Perestroika* by Nancy Ries
ADRIANA PETRYNA
- 238 *Rise and Demise: Comparing World-Systems* by Christopher Chase-Dunn and Thomas D. Hall
DAVID A. SMITH
- 240 *Spaces of Globalization: Reasserting the Power of the Local*, edited by Kevin R. Cox
JANET L. ABU-LUGHOD
- 242 *Consuming the Romantic Utopia: Love and the Cultural Contradictions of Capitalism* by Eva Illouz
SCOTT COLTRANE
- 244 *Work, Ritual, Biography: A Muslim Community in North India* by Deepak Mehta
JAMES ZETKA
- 246 *Time for Life: The Surprising Ways Americans Use Their Time* by John P. Robinson and Geoffrey Godbey
GARY CROSS
- 248 *Out of Place: Homeless Mobilizations, Subcities, and Contested Landscapes* by Talmadge Wright
Something Left to Lose: Personal Relations and Survival among New York's Homeless by Gwendolyn A. Dordick
STEVEN WORDEN
- 250 *The New Middle Class and the Remaking of the Central City* by David Ley
RICHARD HARRIS
- 252 *On My Own: Korean Businesses and Race Relations in America* by In-Jin Yoon
The Korean American Dream: Immigrants and Small Business in New York City by Kyeyoung Park
EDWARD TAEHAN CHANG
- 255 *Just Another Car Factory? Lean Production and Its Discontents* by James Rinehart, Christopher Huxley, and David Robertson
RICHARD FLORIDA

- 256 *Fate and Fortune in Rural China: Social Organisation and Population Behavior in Liaoning, 1774-1873* by James Z. Lee and Cameron D. Campbell
WILLIAM LAVELY
- 258 *How to Succeed in School without Really Learning: The Credentials Race in American Education* by David F. Labaree
CHRISTOPHER J. HURN
- 260 *A History of Young People in the West. Vol. 2, Stormy Evolution to Modern Times*, edited by Giovanni Levi and Jean-Claude Schmitt
SUSAN B. WHITNEY
- 262 *Moral Culture* by Keith Tester
PHILIP SMITH
- 264 *The Magic of the State* by Michael Taussig
SALLY FALK MOORE
- 266 *Authority: Construction and Corrosion* by Bruce Lincoln
JOEL KUIPERS
- 267 *The Ethics of Catholicism and the Consecration of the Intellectual* by André J. Bélanger
LEWIS A. COSER
- 269 *Essays on Religion* by Georg Simmel
LAWRENCE A. SCAFF
- 270 *The Order of Rituals: The Interpretation of Everyday Life* by Hans-Georg Soeffner
R. S. PERINBANAYAGAM
- 273 *The Soul of Development: Biblical Christianity and Economic Transformation in Guatemala* by Amy Sherman
DAVID A. SMILDE
- 275 *Planting Trees in the Developing World: A Sociology of International Organizations* by Steven R. Brechin
FRANK HIRTZ
- 277 *Narrating the Organization: Dramas of Institutional Identity* by Barbara Czarniawska
JOHN W. MOHR
- 279 *Rationalizing Medical Work: Decision-Support Techniques and Medical Practices* by Marc Berg
STEFAN TIMMERMANS

- 281 *Bodies in Protest: Environmental Illness and the Struggle over Medical Knowledge* by Steve Kroll-Smith and H. Hugh Floyd
JUDITH GREEN
- 282 *Community Policing, Chicago Style* by Wesley G. Skogan and Susan M. Hartnett
PETER K. MANNING
- 285 *Just Punishments: Federal Guidelines and Public Views Compared* by Peter H. Rossi and Richard A. Berk
JERRY VAN HOY
- 287 *Overlooking Nazareth: The Ethnography of Exclusion in Galilee* by Dan Rabinowitz
REBECCA L. TORSTICK
- 289 *Der Schweizerische Bauernkrieg von 1653: Politische Sozialgeschichte—Sozialgeschichte eines politischen Ereignisses* by Andreas Suter
MARC R. FORSTER
- 291 *At the Crossroads: Mexican Migration and U.S. Policy* by Frank D. Bean, Rodolfo O. De la Garza, Bryan R. Roberts, and Sidney Weintraub
JORGE DURAND

IN THIS ISSUE

THEODORE P. GERBER is assistant professor of sociology at the University of Oregon. His research focuses on stratification, public attitudes, and political change in Russia, both recently and during the course of the Soviet era. He is principal investigator of the National Science Foundation-funded Survey of Income, Attitudes, and Employment in Russia, for which fieldwork was completed in March 1998. He is using the survey to analyze labor market outcomes since 1990, occupational mobility during the postwar era, and the effects of Soviet-era status on post-Soviet material standing.

MICHAEL HOUT is professor of sociology and director of the Survey Research Center at the University of California, Berkeley. He is one of the six coauthors of *Inequality by Design: Cracking the Bell Curve Myth* (Princeton University Press, 1996). He has written extensively on social mobility, political behavior, and quantitative methods.

PAUL D. MCLEAN is a Harper Instructor in the Social Sciences at the University of Chicago. His research examines the interaction between social networks and identities in the construction and evolution of states, political factions, and markets, particularly in 15th-century Italy and 18th-century Poland.

JAMES D. MONTGOMERY is lecturer in the Interdisciplinary Institute of Management at the London School of Economics. Although trained as an economist, he is probably best described as a mathematical sociologist. His substantive interests include social networks, economic sociology, and the sociology of religion.

MICHAEL T. HANNAN is at Stanford University, where he is professor of organizational behavior and human resources in the Graduate School of Business and professor of sociology. His current research focuses on formalization and integration of theory fragments in organizational ecology, the demography of the corporation, and the evolution of structure and strategy in young, high-technology firms in Silicon Valley.

TOSHIO YAMAGISHI is professor of social psychology in the Faculty of Letters at Hokkaido University, Sapporo. His research interests include social exchange networks, social dilemmas, trust, and social intelligence. He has just published a book, *The Structure of Trust: The Evolutionary Games of Mind and Society* (Tokyo University Press), and is currently conducting a series of cross-societal experiments that provide empirical support for the game theoretic account of individual versus collectivist culture.

KAREN S. COOK is the James B. Duke Professor of Sociology and director of the Laboratory for Experimental Research at Duke University. She has published work on distributive justice, power, and social exchange. Her most recent work is on trust and solidarity in networks of exchange relations.

MOTOKI WATABE is a doctoral student in the Sociology Department at the University of California, Los Angeles. His research interests include social dilemmas and social networks. He is currently interested in the role of trust for the network formation in social dilemma situations.

MICHAEL J. LOVAGLIA is associate professor of sociology at the University of Iowa. Current research interests include power and status in exchange networks and the relationship between emotion and social structure. He edits *Current Research in Social Psychology*, an electronic journal, currently available at <http://www.uiowa.edu/~grpproc>.

JEFFREY W. LUCAS is a Ph.D. candidate in the Sociology Department at the University of Iowa. His areas of interest include social psychology and organizations. In addition to studying power in exchange networks, his current research investigates the impact of job titles on satisfaction, commitment, and performance in organizational work.

JEFFREY A. HOUSER is an assistant professor at Bowling Green State University. His research interests include status processes, the social psychology of disability, and emotion. He has recently published articles on the impact of physical disability on social status, and the theoretical foundations of the Americans with Disabilities Act.

SHANE R. THYE is a postdoctoral research fellow in the Department of Organizational Behavior at the School of Industrial and Labor Relations, Cornell University. He is currently investigating power and commitment in exchange relations, the spread of status value in networks, emotional solidarity, and learning models of cooperation in social dilemmas. He will join the Sociology Department at the University of South Carolina in fall 1998.

BARRY MARKOVSKY, professor of sociology at the University of Iowa, currently is the program director for sociology at the National Science Foundation. His research interests are in the areas of group processes, social psychology, social networks, and issues in theory construction and analysis.

Information for Contributors

Editorial Procedures: All papers deemed appropriate for the *AJS* are sent out anonymously to readers. The editors rely heavily on the judgments of those readers, although they are not bound by them. To protect anonymity, only the title should appear on the manuscript. Attach a cover page with the title of the manuscript, your name, and your affiliation. Manuscripts are accepted for publication subject to the understanding that the *AJS* has the right of first publication. Submission of a manuscript to the *AJS* is taken to indicate the author's commitment to publish in this *Journal*. No paper known to be under consideration by any other journal will be reviewed by the *AJS*. In addition, the *AJS* is not a suitable vehicle for materials published elsewhere, in whole or in part. Upon publication, all rights, including subsidiary rights, are owned by the University of Chicago Press. The author retains the right to use his or her article without charge in any book of which he or she is the author or editor after it has appeared in the *Journal*.

Manuscript Acceptance Policy: While it is our policy to require the assignment of copyright on most journal articles and review essays, we do not usually request assignment of copyright for other contributions. Although the copyright to such a contribution may remain with the author, it is understood that, in return for publication, the *Journal* has the nonexclusive right to publish the contribution and the continuing right, without limit, to include the contribution as part of any reprinting of the issue and/or volume of the *Journal* in which the contribution first appeared by any means and in any format, including computer-assisted storage and readout, in which the issue and/or volume may be reproduced by the publisher or by its licensed agencies.

Preparation of Copy

1. Type *all* copy—including indented matter, footnotes, tables, and references—*double spaced*, allowing generous margins at the top, bottom, and sides of page. Please do not break words at ends of lines or justify right-hand margins. Do not use proportional spacing fonts. Indicate italics by underlining only. Number all footnotes, beginning with the acknowledgment note as number 1. *Provide a word count on the cover page of the manuscript.*

2. Begin each table on a separate page. Refer to each table in numerical order in the text. Tables should not contain vertical lines or more than 20 two-digit columns or the equivalent.

3. Clarify all mathematical symbols (e.g., Greek letters) with words in the margins of the manuscript.

4. Original or glossy prints of professionally drawn figures may be required if manuscript is accepted.

5. Include a brief abstract (not more than 100 words) summarizing the findings.

6. Submit *five copies* of the manuscript and abstract on white paper. Manuscripts should be addressed to the Editor of the AMERICAN JOURNAL OF SOCIOLOGY, 5835 South Kimbark, Chicago, Illinois 60637. *Manuscripts are not returned to authors.* Enclose a self-addressed stamped postcard to obtain acknowledgment of manuscript receipt. No other acknowledgment will be sent.

7. Each article submitted to the *AJS* must be accompanied by a check or money order for \$15.00, payable to the University of Chicago Press in U.S. currency or

its equivalent. Papers will not be reviewed until this fee has been paid. The submission fee will be waived for papers solely by student authors. Submissions from students must be accompanied by a faculty member's certification of their student status. Citizens of countries with restrictions on the export of U.S. dollars may request waiver of the submission fee.

Format of References in Text: All references to monographs, articles, and statistical sources are to be identified at the appropriate point in the text by last name of author, year of publication, and pagination where appropriate, all within parentheses. Footnotes are to be used only for substantive observations. Specify subsequent citations of the same source in the same way as the first one; do not use "ibid.," "op. cit.," or "loc. cit."

1. When author's name is in the text: Duncan (1959). When author's name is not in text: (Gouldner 1963).

2. Pagination follows year of publication: (Lipset 1964, pp. 61-65).

3. For more than three authors, use "et al." For institutional authorship, supply minimum identification from the beginning of the complete citation: (U.S. Bureau of the Census 1963, p. 117).

4. With more than one reference to an author in the same year, distinguish them by use of letters (*a*, *b*) attached to the year of publication: (1965*a*).

5. Enclose a series of references with a single pair of parentheses, separated by semicolons: (Duncan 1959; Gouldner 1963; Lipset 1964, pp. 61-65).

6. Format for references to machine-readable data files is

Davis, James Allan. 1978. GSS (General Social Survey). NORC ed. Chicago: NORC. Distributed by Roper Public Opinion Research Center, New Haven, Conn.

Format of References: List all items, including machine-readable data files (MRDF), alphabetically by author (providing the full list of multiple authors) and, within author(s), by year of publication, in an appendix titled "REFERENCES." For examples, see text of articles in this issue.

(Rev. 3/98)

CONTENTS

- 293 The Meaning of Data: Open and Closed Evidential Cultures in the Search for Gravitational Waves
H. M. COLLINS

- 339 Making Use of the Past: Time Periods as Cases to Compare and as Sequences of Problem Solving
JEFFREY HAYDU

- 372 Defended Neighborhoods, Integration, and Racially Motivated Crime
DONALD P. GREEN, DARA Z. STROLOVITCH, AND JANELLE S. WONG

- 404 Engineering Growth: Business Group Structure and Firm Performance in China's Transition Economy
LISA A. KEISTER

- 441 The Strength of a Weak State: The Rights Revolution and the Rise of Human Resources Management Divisions
FRANK DOBBIN AND JOHN R. SUTTON

- 477 High-Risk Collective Action: Defending Human Rights in Chile, Uruguay, and Argentina
MARA LOVEMAN

Book Reviews

- 526 *Of Revelation and Revolution*. Vol. 2, *The Dialectics of Modernity on a South African Frontier* by John L. Comaroff and Jean Comaroff
DAVID CHIDESTER

- 528 *Struggle for the Spirit: Religious Transformation and Popular Culture in Brazil and Latin America* by David Lehmann
JOHN BURDICK

- 530 *Disquiet in the Land: Cultural Conflict in American Mennonite Communities* by Fred Kniss
COURTNEY BENDER
- 532 *Reinventing American Protestantism: Christianity in the New Millennium* by Donald E. Miller
NANCY L. EIESLAND
- 534 *God's Daughters: Evangelical Women and the Power of Submission* by R. Marie Griffith
SUSAN D. ROSE
- 537 *Splintered Sisterhood: Gender and Class in the Campaign against Woman Suffrage* by Susan E. Marshall
KATHLEEN M. BLEE
- 539 *Changing the Wor(l)d: Discourse, Politics, and the Feminist Movement* by Stacey Young
KATIE KING
- 541 *Beyond Hierarchy: Gender, Sexuality and Social Economy* by Sarah Oerton
JOYCE ROTHSCILD
- 543 *Crisis and Political Beliefs: The Case of the Colt Firearms Strike* by Marc Lendler
JUDITH STEPAN-NORRIS
- 544 *Don't Burn It Here: Grassroots Challenges to Trash Incinerators* by Edward J. Walsh, Rex Warland, and D. Clayton Smith
THOMAS E. SHRIVER
- 546 *The End of the Peasantry: The Rural Labor Movement in Northeast Brazil, 1961-1988* by Anthony W. Pereira
NANCY D. LAPP
- 548 *Peasant Protest and Social Change in Colonial Korea* by Gi-Wook Shin
HAGEN KOO
- 550 *Taking on Goliath: The Emergence of a New Left Party and the Struggle for Democracy in Mexico* by Kathleen Bruhn
KEVIN J. MIDDLEBROOK
- 552 *Urban Protest in Seventeenth-Century France: The Culture of Re-tribution* by William Beik
JOHN MARKOFF

- 554 *Radical Artisans in England and France, 1830–1870* by Iowerth Prothero
RICHARD BIERNACKI
- 556 *Return to the Western World: Cultural and Political Perspectives on the Estonian Post-Communist Transition*, edited by Marju Lauristin, Peeter Vilhalemm, Karl Erik Rosengren, and Lennart Weibull
JOAN LÖFGREN
- 558 *Money Politics in the New Europe: Britain, France and the Single Financial Market* by Daphne Josselin
DAVID KNOKE
- 560 *Modernisation and Postmodernization: Cultural, Economic, and Political Change in 43 Societies* by Ronald Inglehart
AVNER OFFER
- 562 *From Subject to Citizen: Australian Citizenship in the Twentieth Century* by Alastair Davidson
JEFF SPINNER-HALEV
- 564 *Molding Japanese Minds: The State in Everyday Life* by Sheldon Garon
TETSUYA KATAOKA
- 566 *The Rule of Law in the Arab World: Courts in Egypt and the Gulf* by Nathan J. Brown
LAWRENCE ROSEN
- 567 *Jews in Germany after the Holocaust: Memory, Identity, and Jewish-German Relations* by Lynn Rapaport
ANDREI S. MARKOVITS
- 570 *Internationalism and Its Betrayal* by Micheline R. Ishay
ROMAN SZPORLUK
- 571 *Social Darwinism in European and American Thought, 1860–1945: Nature as Model and Nature as Threat* by Mike Hawkins
LEONARD MOORE
- 573 *Emancipation and Illusion: Rationality and Gender in Habermas's Theory of Modernity* by Marie Fleming
JODI DEAN
- 575 *The New Public: Professional Communication and the Means of Social Influence* by Leon H. Mayhew
NICHOLAS GARNHAM

- 577 *World in Fragments: Writings on Politics, Society, Psychoanalysis, and the Imagination* by Cornelius Castoriadis
THOMAS M. KEMPLE
- 579 *Back to Reality? Social Experience and Cultural Studies*, edited by Angela McRobbie
HERMAN SANDERS GRAY
- 581 *Anthropological Demography: Toward a New Synthesis*, edited by David I. Kertzer and Tom Fricke
ALAKA M. BASU
- 582 *Auto/Ethnography: Rewriting the Self and the Social*, edited by Deborah Reed-Danahay
After Writing Culture: Epistemology and Praxis in Contemporary Anthropology, edited by Allison James, Jenny Hockey, and Andrew Dawson
GEORGE E. MARCUS
- 585 *Redeeming Laughter: The Comic Dimension of Human Experience* by Peter L. Berger
MAHADEV APTE
- 587 *Misery and Company: Sympathy in Everyday Life* by Candace Clark
JACK KATZ
- 589 *Risk and Misfortune: A Social Construction of Accidents* by Judith Green
JOSEPH GUSFIELD
- 591 *Making Crime Pay: Law and Order in Contemporary American Politics* by Katherine Beckett
DAVID JACOBS
- 592 *Penitentiaries, Reformatories, and Chain Gangs: Social Theory and the History of Punishment in Nineteenth-Century America* by Mark Colvin
THOMAS F. COURTLESS
- 594 *Causalities of Community Disorder: Women's Careers in Violent Crime* by Deborah R. Baskin and Ira B. Sommers
PAT CARLEN

IN THIS ISSUE

H. M. COLLINS is Distinguished Research Professor of Sociology and director of the Centre for the Study of Knowledge, Expertise, and Science (KES) at Cardiff University. He is the 1997 recipient of the J. D. Bernal Award of the Society for Social Studies of Science and is currently working on a longitudinal study of the history and sociology of gravitational wave physics.

JEFFREY HAYDU is the author of *Between Craft and Class* (1988) and *Making American Industry Safe for Democracy* (1997). He is currently studying employer class formation and ideologies of workplace justice in the late 19th century.

DONALD P. GREEN is professor of political science and director of the Institution for Social and Policy Studies at Yale University.

DARA Z. STROLOVICH is a doctoral candidate in the Department of Political Science at Yale University. Her research interests include public opinion, political representation, and the politics of race and gender.

JANELLE S. WONG is a doctoral candidate in the Department of Political Science at Yale University. Her research interests include racial and ethnic politics and public opinion.

LISA A. KEISTER is assistant professor of sociology at the University of North Carolina, Chapel Hill. She is currently working on a book examining the structure and economic impact of Chinese business groups and related papers on the role of the groups in the emergence of a financial market in China. She is also working on a study of family wealth accumulation and distribution in the United States.

FRANK DOBBIN is associate professor of sociology at Princeton University. He is spending the year, with a Russell Sage Foundation fellowship, working on a book on changing employer response to affirmative action law. His book, *Forging Industrial Policy*, received the 1996 Max Weber Award from the American Sociological Association's section on organizations and occupations.

JOHN R. SUTTON is professor of sociology at the University of California, Santa Barbara. His paper in this issue is the latest in a series of reports on organizations and their legal environments that he has coauthored with Frank Dobbin, W. Richard Scott, and John W. Meyer. He is currently studying the institutional ecology of imprisonment in modern Western democracies.

MARA LOVEMAN is a doctoral student in the Department of Sociology at the University of California, Los Angeles, and a Mellon Fellow in Latin American sociology. Her research interests include social theory, comparative historical methodology, social movements, political sociology, and "race," ethnicity, and nationhood in comparative historical perspective.

Information for Contributors

Editorial Procedures: All papers deemed appropriate for the *AJS* are sent out anonymously to readers. The editors rely heavily on the judgments of those readers, although they are not bound by them. To protect anonymity, only the title should appear on the manuscript. Attach a cover page with the title of the manuscript, your name, and your affiliation. Manuscripts are accepted for publication subject to the understanding that the *AJS* has the right of first publication. Submission of a manuscript to the *AJS* is taken to indicate the author's commitment to publish in this *Journal*. No paper known to be under consideration by any other journal will be reviewed by the *AJS*. In addition, the *AJS* is not a suitable vehicle for materials published elsewhere, in whole or in part. Upon publication, all rights, including subsidiary rights, are owned by the University of Chicago Press. The author retains the right to use his or her article without charge in any book of which he or she is the author or editor after it has appeared in the *Journal*.

Manuscript Acceptance Policy: While it is our policy to require the assignment of copyright on most journal articles and review essays, we do not usually request assignment of copyright for other contributions. Although the copyright to such a contribution may remain with the author, it is understood that, in return for publication, the *Journal* has the nonexclusive right to publish the contribution and the continuing right, without limit, to include the contribution as part of any reprinting of the issue and/or volume of the *Journal* in which the contribution first appeared by any means and in any format, including computer-assisted storage and readout, in which the issue and/or volume may be reproduced by the publisher or by its licensed agencies.

Preparation of Copy

1. Type *all* copy—including indented matter, footnotes, tables, and references—*double spaced*, allowing generous margins at the top, bottom, and sides of page. Please do not break words at ends of lines or justify right-hand margins. Do not use proportional spacing fonts. Indicate italics by underlining only. Number all footnotes, beginning with the acknowledgment note as number 1. *Provide a word count on the cover page of the manuscript.*

2. Begin each table on a separate page. Refer to each table in numerical order in the text. Tables should not contain vertical lines or more than 20 two-digit columns or the equivalent.

3. Clarify all mathematical symbols (e.g., Greek letters) with words in the margins of the manuscript.

4. Original or glossy prints of professionally drawn figures may be required if manuscript is accepted.

5. Include a brief abstract (not more than 100 words) summarizing the findings.

6. Submit *five copies* of the manuscript and abstract on white paper. Manuscripts should be addressed to the Editor of the AMERICAN JOURNAL OF SOCIOLOGY, 5835 South Kimbark, Chicago, Illinois 60637. *Manuscripts are not returned to authors.* Enclose a self-addressed stamped postcard to obtain acknowledgment of manuscript receipt. No other acknowledgment will be sent.

7. Each article submitted to the *AJS* must be accompanied by a check or money order for \$15.00, payable to the University of Chicago Press in U.S. currency or

its equivalent. Papers will not be reviewed until this fee has been paid. The submission fee will be waived for papers solely by student authors. Submissions from students must be accompanied by a faculty member's certification of their student status. Citizens of countries with restrictions on the export of U.S. dollars may request waiver of the submission fee.

Format of References in Text: All references to monographs, articles, and statistical sources are to be identified at the appropriate point in the text by last name of author, year of publication, and pagination where appropriate, all within parentheses. Footnotes are to be used only for substantive observations. Specify subsequent citations of the same source in the same way as the first one; do not use "ibid.," "op. cit.," or "loc. cit."

1. When author's name is in the text: Duncan (1959). When author's name is not in text: (Gouldner 1963).

2. Pagination follows year of publication: (Lipset 1964, pp. 61-65).

3. For more than three authors, use "et al." For institutional authorship, supply minimum identification from the beginning of the complete citation: (U.S. Bureau of the Census 1963, p. 117).

4. With more than one reference to an author in the same year, distinguish them by use of letters (a, b) attached to the year of publication: (1965a).

5. Enclose a series of references with a single pair of parentheses, separated by semicolons: (Duncan 1959; Gouldner 1963; Lipset 1964, pp. 61-65).

6. Format for references to machine-readable data files is

Davis, James Allan. 1978. GSS (General Social Survey). NORC ed. Chicago: NORC. Distributed by Roper Public Opinion Research Center, New Haven, Conn.

Format of References: List all items, including machine-readable data files (MRDF), alphabetically by author (providing the full list of multiple authors) and, within author(s), by year of publication, in an appendix titled "REFERENCES." For examples, see text of articles in this issue.

(Rev. 3/98)

CONTENTS

- 597 A Cultural Field in the Making: Gastronomy in 19th-Century France
PRISCILLA PARKHURST FERGUSON

- 642 Linking Action to Social Structure within a System: Social Capital within and between Subgroups
KENNETH A. FRANK AND JEFFREY Y. YASUMOTO

- 687 Origins, Destinations, and Association in Occupational Mobility
MICHAEL E. SOBEL, MARK P. BECKER, AND SUSAN M. MINICK

Symposium on Historical Sociology and Rational Choice Theory

- 722 "We're No Angels": Realism, Rational Choice, and Relationality in Social Science
MARGARET R. SOMERS

- 785 The Debate on Historical Sociology: Rational Choice Theory and Its Critics
EDGAR KISER AND MICHAEL HECHTER

- 817 Limitations of Rational Choice Theory
RAYMOND BOUDON

- 829 Initial Conditions, General Laws, Path Dependence, and Explanation in Historical Sociology
JACK A. GOLDSTONE

- 846 Explanation in Historical Sociology: Narrative, General Theory, and Historically Specific Theory
CRAIG CALHOUN

Commentary and Debate

- 872 Rawls, Durkheim, and Causality: A Critical Discussion
WARREN SCHMAUS

- 887 Durkheim's Challenge to Philosophy: Human Reason Explained
as a Product of Enacted Social Practice
ANNE WARFIELD RAWLS

Review Essay

- 902 Chasing Alfred Chandler
NEIL FLIGSTEIN

Book Reviews

- 906 *Poverty, Ethnicity, and Violent Crime* by James F. Short, Jr.
ROBERT J. BURSİK, JR.
- 908 *Violence and Childhood in the Inner City*, edited by Joan McCord
DARNELL F. HAWKINS
- 909 *Violence and Truth: A Rajasthani Kingdom Confronts Colonial
Authority* by Denis Vidal
FRANK DE ZWART
- 911 *Racism in a Racial Democracy: The Maintenance of White Su-
premacy in Brazil* by France Winddance Twine
EVANDRO CAMARA
- 913 *Nationalising Blackness: Afrocubanismo and Artistic Revolution
in Havana, 1920-1940* by Robin D. Moore
LAUREN DERBY
- 916 *Inside the Revolution: Everyday Life in Socialist Cuba* by Mona
Rosendahl
SHERYL L. LUTJENS
- 918 *Salt of the Earth: The Political Origins of Peasant Protest and
Communist Revolution in China* by Ralph A. Thaxton, Jr.
R. DAVID ARKUSH
- 920 *Movement Genesis: Social Movement Theory and the 1980s West
German Peace Movement* by Steve Breyman
DAVID S. MEYER
- 922 *The Social Psychology of Protest* by Bert Klandermans
JAMES M. JASPER
- 923 *Stalinism and Nazism: Dictatorships in Comparison*, edited by Ian
Kershaw and Moshe Lewin
DANIEL CHIROT

- 925 *The Policing of Politics in the Twentieth Century: Historical Perspectives*, edited by Mark Mazower
LAURA KALMANOWIECKI
- 927 *Settling Accounts: Violence, Justice, and Accountability in Post-socialist Europe* by John Borneman
STANLEY COHEN
- 929 *The Great Surprise of the Small Transformation: The Demise of Communism and the Rise of the Private Sector in Hungary* by Ákos Róna-Tas
TONY LEVITAS
- 931 *The Testimony of Lives: Narrative and Memory in Post-Soviet Latvia* by Vieda Skultans
MIKK TITMA
- 932 *Power and the Sacred in Revolutionary Russia: Religious Activists in the Village* by Glennys Young
RICHARD L. HERNANDEZ
- 934 *Sacred Tensions: Modernity and Religious Transformation in Malaysia* by Raymond L. M. Lee and Susan E. Ackerman
ROBERT W. HEFNER
- 936 *Millennium, Messiahs, and Mayhem: Contemporary Apocalyptic Movements*, edited by Thomas Robbins and Susan J. Palmer
JAMES T. RICHARDSON
- 939 *Contemporary American Religion: An Ethnographic Reader*, edited by Penny Edgell Becker and Nancy L. Eiesland
DANIEL V. A. OLSON
- 941 *Cultural Wars in American Politics: Critical Reviews of a Popular Myth*, edited by Rhys H. Williams
WADE CLARK ROOF
- 943 *Making American Industry Safe for Democracy: Comparative Perspectives on the State and Employee Representation in the Era of World War I* by Jeffrey Haydu
KATHLEEN THELEN
- 945 *Citizen Politics in Post-Industrial Societies*, edited by Terry Nichols Clark and Michael Rempel
JANE JUNN

- 946 *Modern Manors: Welfare Capitalism since the New Deal* by Sanford M. Jacoby
CHRISTOPHER HOWARD
- 948 *The Hidden Welfare State: Tax Expenditures and Social Policy in the United States* by Christopher Howard
EDWIN AMENTA
- 950 *Labour Market Regimes and Patterns of Flexibility: A Sweden-Canada Comparison* by Axel van den Berg, Bengt Furåker, and Leif Johansson
MARY RUGGIE
- 952 *Economic Ideology and Japanese Industrial Policy: Developmentalism from 1931 to 1965* by Bai Gao
J. MARK RAMSEYER
- 954 *Along the Domestic-Foreign Frontier: Exploring Governance in a Turbulent World* by James N. Rosenau
TONY MCGREW
- 956 *The Challenge of Interracial Unionism: Alabama Coal Miners, 1878-1921* by Daniel Letwin
JOHN LIE
- 957 *Paths to Success: Beating the Odds in American Society* by Charles C. Harrington and Susan K. Boardman
JENNIFER L. HOCHSCHILD
- 959 *Against the Odds? Social Class and Social Justice in Industrial Societies* by Gordon Marshall, Adam Swift, and Stephen Roberts
CHARLES E. HURST
- 960 *Capital Culture: Gender at Work in the City* by Linda McDowell
AMY S. WHARTON
- 963 *Liberty, Equality, and Justice: Civil Rights, Women's Rights, and the Regulation of Business, 1865-1932* by Ross Evans Paulson
EILEEN BORIS
- 965 *Creating Socialist Women in Japan: Gender, Labour and Activism, 1900-1937* by Vera Mackie
SHARON SIEVERS
- 967 *Real Knockouts: The Physical Feminism of Women's Self-Defense* by Martha McCaughey
BONNIE H. ERICKSON

- 968 *Media and Political Conflict: News from the Middle East* by Gadi Wolfsfeld
DARNELL M. HUNT
- 970 *Music and Cultural Theory* by John Shepherd and Peter Wicke
ROBERT W. WITKIN
- 972 *The Rules of Art: Genesis and Structure of the Literary Field* by Pierre Bourdieu
WENDY GRISWOLD
- 975 *The Business of Children's Entertainment* by Norma Odom Pecora
HEATHER HENDERSHOT
- 977 *Community Organizing for Urban School Reform* by Dennis Shirley
JEAN ANYON
- 978 *Against the Multicultural Agenda: A Critical Thinking Alternative*
by Yehudi O. Webster
JAMES A. BANKS
- 979 *Children, Schools, and Inequality*, edited by Doris R. Entwisle,
Karl L. Alexander, and Linda Steffel Olson
BARBARA HEYNS
- 981 *Motivation and Delinquency*, vol. 44 of the Nebraska Symposium
on Motivation edited by D. Wayne Osgood
SIMON I. SINGER

IN THIS ISSUE

PRISCILLA PARKHURST FERGUSON is professor of sociology at Columbia University where she also is a member of the Interdisciplinary Committee on Comparative Literature. Her books include *Literary France: The Making of a Culture* (1987) and *Paris as Revolution: Reading the Nineteenth-Century City* (1994), and she is currently at work on *Eating Orders*, a sociological exploration of cuisine as a vehicle of cultural identity.

KENNETH A. FRANK is assistant professor in the Department of Counseling, Educational Psychology, and Special Education within the College of Education at Michigan State University. His methodological interests include multilevel models and analyses of social network data including clustering and modeling. His substantive interests include organizational theory, specifically the study of schools as organizations.

JEFFREY Y. YASUMOTO is a doctoral candidate in sociology at the University of Chicago. His major areas of interest include formal organizations, social networks, and sociology of education. His current work includes coauthoring a book on the impact of teacher networks on student achievement, dissertation work on the social organization of students, and an article on pricing in financial markets.

MICHAEL E. SOBEL, professor of sociology and applied mathematics, is currently working on the topics of causal inference and modeling categorical data. He coedited the *Handbook of Statistical Modeling for the Social and Behavioral Sciences* (Plenum Press, 1995) and is currently coeditor of *Sociological Methodology*.

MARK P. BECKER is professor of biostatistics and associate dean for academic affairs in the School of Public Health and a faculty associate in the Institute for Social Research, all at the University of Michigan.

SUSAN M. MINICK is assistant professor in the Department of Medicine, Division of Biostatistics, at Indiana University, Indianapolis. Her research interests include the development of methods for the analysis of correlated categorical outcomes in areas such as rater agreement and social mobility.

MARGARET R. SOMERS is A. Bartlett Giamatti Faculty Fellow, Institute for the Humanities, and associate professor of sociology and history at the University of Michigan. Her research includes European politics, citizenship rights, and state formation; historical institutionalism and economic sociology; and epistemology and science. She is the author of *Citizenship and the Public Sphere: Institutions, Associations, Rights* (Cambridge University Press, in press).

EDGAR KISER is associate professor of sociology at the University of Washington. He has published articles in sociology, political science, and economics journals on topics including the determinants of war, the development and decline of voting institutions, and the evolution of tax systems. In addition to writing a book tentatively entitled *The Rise of the Bureaucratic State*, he is currently working on the growth of the state in

early modern France (with April Eaton), the anomalous early bureaucratization in Qin China (with Yong Cai), the relationship between war, revolution, and state making in 19th-century Europe (with Josh Kane), and the determinants of tax rates on business in contemporary states (with Aaron Laing).

MICHAEL HECHTER is professor of sociology at the University of Arizona. His books include *Internal Colonialism: The Celtic Fringe in British National Development* (2d ed., 1998), *Principles of Group Solidarity* (1987), and the forthcoming *Containing Nationalism*.

RAYMOND BOUDON is professor at the University of Paris-Sorbonne, member of the American Academy of Arts and Sciences and of the Académie de Sciences Morales et Politiques. He is presently working on the sociological theory of values.

JACK A. GOLDSTONE is professor of sociology at the University of California, Davis. His latest work is entitled *How Societies Change: Revolutions, Social Movements, and the Reproduction of Societies* (Harvard University Press, in press), and he is editor in chief of the *Encyclopedia of Political Revolutions* (Congressional Quarterly Books, 1998).

CRAIG CALHOUN is professor of sociology at New York University.

WARREN SCHMAUS wrote the article featured here while he was a fellow at the Center for the Philosophy of Science at the University of Pittsburgh. He is the author of *Durkheim's Philosophy of Science and the Sociology of Knowledge* (University of Chicago Press, 1994) and is currently professor of philosophy at Illinois Institute of Technology.

ANNE WARFIELD RAWLS is associate professor of sociology at Wayne State University. Her writings bring together issues in philosophy, language studies, and social theory in an effort to address sociologically essential issues of justice and equality.

NEIL FLIGSTEIN is professor of sociology at the University of California, Berkeley.

Erratum

In the July issue (*AJS* 104, p. 240), a miscopied transcription led to an incorrect citation in Janet Abu-Lughod's review of *Spaces of Globalization*. The line should have said "Fröbel, Heinrichs, and Kreye's *New International Division of Labor*." We regret the error.

Information for Contributors

Editorial Procedures: All papers deemed appropriate for the *AJS* are sent out anonymously to readers. The editors rely heavily on the judgments of those readers, although they are not bound by them. To protect anonymity, only the title should appear on the manuscript. Attach a cover page with the title of the manuscript, your name, and your affiliation. Manuscripts are accepted for publication subject to the understanding that the *AJS* has the right of first publication. Submission of a manuscript to the *AJS* is taken to indicate the author's commitment to publish in this *Journal*. No paper known to be under consideration by any other journal will be reviewed by the *AJS*. In addition, the *AJS* is not a suitable vehicle for materials published elsewhere, in whole or in part. Upon publication, all rights, including subsidiary rights, are owned by the University of Chicago Press. The author retains the right to use his or her article without charge in any book of which he or she is the author or editor after it has appeared in the *Journal*.

Manuscript Acceptance Policy: While it is our policy to require the assignment of copyright on most journal articles and review essays, we do not usually request assignment of copyright for other contributions. Although the copyright to such a contribution may remain with the author, it is understood that, in return for publication, the *Journal* has the nonexclusive right to publish the contribution and the continuing right, without limit, to include the contribution as part of any reprinting of the issue and/or volume of the *Journal* in which the contribution first appeared by any means and in any format, including computer-assisted storage and readout, in which the issue and/or volume may be reproduced by the publisher or by its licensed agencies.

Preparation of Copy

1. Type *all* copy—including indented matter, footnotes, tables, and references—*double spaced*, allowing generous margins at the top, bottom, and sides of page. Please do not break words at ends of lines or justify right-hand margins. Do not use proportional spacing fonts. Indicate italics by underlining only. Number all footnotes, beginning with the acknowledgment note as number 1. *Provide a word count on the cover page of the manuscript.*

2. Begin each table on a separate page. Refer to each table in numerical order in the text. Tables should not contain vertical lines or more than 20 two-digit columns or the equivalent.

3. Clarify all mathematical symbols (e.g., Greek letters) with words in the margins of the manuscript.

4. Original or glossy prints of professionally drawn figures may be required if manuscript is accepted.

5. Include a brief abstract (not more than 100 words) summarizing the findings.

6. Submit *five copies* of the manuscript and abstract on white paper. Manuscripts should be addressed to the Editor of the AMERICAN JOURNAL OF SOCIOLOGY, 5835 South Kimbark, Chicago, Illinois 60637. *Manuscripts are not returned to authors.* Enclose a self-addressed stamped postcard to obtain acknowledgment of manuscript receipt. No other acknowledgment will be sent.

7. Each article submitted to the *AJS* must be accompanied by a check or money order for \$15.00, payable to the University of Chicago Press in U.S. currency or

its equivalent. Papers will not be reviewed until this fee has been paid. The submission fee will be waived for papers solely by student authors. Submissions from students must be accompanied by a faculty member's certification of their student status. Citizens of countries with restrictions on the export of U.S. dollars may request waiver of the submission fee.

Format of References in Text: All references to monographs, articles, and statistical sources are to be identified at the appropriate point in the text by last name of author, year of publication, and pagination where appropriate, all within parentheses. Footnotes are to be used only for substantive observations. Specify subsequent citations of the same source in the same way as the first one; do not use "ibid.," "op. cit.," or "loc. cit."

1. When author's name is in the text: Duncan (1959). When author's name is not in text: (Gouldner 1963).

2. Pagination follows year of publication: (Lipset 1964, pp. 61-65).

3. For more than three authors, use "et al." For institutional authorship, supply minimum identification from the beginning of the complete citation. (U.S. Bureau of the Census 1963, p. 117).

4. With more than one reference to an author in the same year, distinguish them by use of letters (*a*, *b*) attached to the year of publication: (1965*a*).

5. Enclose a series of references with a single pair of parentheses, separated by semicolons: (Duncan 1959; Gouldner 1963; Lipset 1964, pp. 61-65).

6. Format for references to machine-readable data files is

Davis, James Allan. 1978. GSS (General Social Survey). NORC ed. Chicago: NORC. Distributed by Roper Public Opinion Research Center, New Haven, Conn.

Format of References: List all items, including machine-readable data files (MRDF), alphabetically by author (providing the full list of multiple authors) and, within author(s), by year of publication, in an appendix titled "REFERENCES." For examples, see text of articles in this issue.

(Rev. 3/98)

More Shock than Therapy: Market Transition, Employment, and Income in Russia, 1991–1995¹

Theodore P. Gerber
University of Oregon

Michael Hout
University of California, Berkeley

Sixteen predictions from market transition theory are assessed using survey data on employment, earnings, and income in Russia, during the first five years of market reform. Although the private sector has grown, self-employment is still rare. Incomes are down, and unemployment is up. A distended income distribution reflects unprecedented income inequality. Distinctive features of late Soviet-era stratification persist: low returns to education, a gender gap in earnings, and low earnings among professionals. The Russian market transition offers more opportunity in trade, consumer services, and speculation and less in manufacturing than do other emerging markets. This corresponds to “merchant capitalism” and contradicts the predictions of market transition theory.

Everything the Communists told us about communism was a lie. Everything they told us about capitalism turns out to be true. (Popular Russian joke, circa 1996)

INTRODUCTION

Most state socialist countries in Europe and Asia adopted sweeping market reforms during the 1980s and 1990s. Victor Nee took note of how these reforms usually seek to weaken the tie between the state and the economy

¹ Some of the results in this article were presented at meetings of the Research Committee on Social Stratification (RC28) of the International Sociological Association, Ann Arbor, Michigan, August 1996; the American Sociological Association, New York City, August 1996; and RC28, Tel Aviv, May 1997. The authors would like to thank Nancy Tuma, Neil Fligstein, Yu Xie, Ylon Cohen, Jane Zavitsa, and the *AJS* reviewers for helpful criticisms on an earlier draft, Ludmila Khakhulina and Larisa Koseova of the All-Russian Center for the Study of Public Opinion for assistance with the data and discussions of the issues presented in this article, and Edward Akhmetshin for research assistance. This article was written while Hout was a visiting scholar at the Russell Sage Foundation; we thank the foundation for financial assistance and

and proposed a broad theory of how market transition alters social stratification (Nee 1989, 1991, 1996; Nee and Cao 1996; Nee and Matthews 1996). Generalizing from their research in China, Nee and his associates developed a market transition theory that emphasizes the role of private firms and individuals in making markets within a (formerly) socialist redistributive system. The key dynamic in market transition theory occurs as the state withdraws and alternative paths to economic success open. Key propositions of the theory include the following: (1) market opportunities enhance returns to both education and entrepreneurial initiative; and (2) market transition diminishes the rentlike advantages political elites derive from their redistributive power. On balance, inequality due to rising returns to human capital and initiative is offset by decreased inequality on the basis of political capital.

Market transition theory is controversial. Some scholars see less individual opportunity in China. Communist Party members continue to earn more than others, as do managers who hold redistributive power (Bian and Logan 1996; Xie and Hannum 1996).² Persistent effects do not imply that the Chinese economic system is static. As Parish and Michelson (1996, p. 1056) conclude, "In the Chinese countryside there is both considerable market and politically induced flux, with many individuals seeking fame and fortune through nonadministrative channels and new types of administrators rising to the top; at the same time local administrators have considerable opportunity, *particularly in more marketized and corporatized areas*" (our emphasis). These findings are consistent with Walder's (1995b) observation that reform can often provide incentives within the sphere of public ownership, at times more effectively than relying on privatization can.

Outside of China, market transitions have varied in their pace, strategy, results, and political context (see Brada 1996; Murrell 1996). Generalizing from China's experience carries some intellectual risks. Róna-Tas (1994) concludes that the transition in Hungary offers few opportunities for individuals who lack organizational resources; he characterizes the changes as "power conversion" or "technocratic continuity." Stark (1996) notes that public and privatized managers equally seek to manage risk, resulting in networks and organizational forms he calls "recombinant property forms."

The course of economic change in Russia is, without doubt, relevant to these debates. Considering its vast size, economic and military power,

its myriad other benefits. Direct correspondence to Theodore P. Gerber, Department of Sociology, 1291, University of Oregon, Eugene, Oregon 97401. E-mail: tgerber@oregon.uoregon.edu

² As Walder (1996, p. 1068) points out, Nee's data also show the persistence of cadres' income advantage.

and historical status as the birthplace of state socialism, general conclusions about the impact of market reforms on the magnitude and mechanisms of socioeconomic inequalities must take account of Russia. But, to date, discussions of market transition theory have not included systematic analysis of the Russian case.

When the federal government implemented radical economic reforms in January 1992, President Boris Yeltsin promised to raise Russians' standard of living. Experts—both foreign and domestic—predicted that “shock therapy” would displace people for a short time but lead to growth before long (Sachs 1992). Press reports of declining living standards and the electoral success of a resurgent Communist Party in 1996 cast doubt on the success of market reforms. Thus, employment and income trends in Russia are basic social indicators valued as much as political harbingers as data for testing theory.

Using five surveys spanning from February 1991 to January 1996, we analyze trends in employment, earnings, and the mechanisms of income stratification over the course of the Russian market transition. Our results show that there has been more shock than therapy in post-Soviet Russia. Although the private sector has grown, self-employment is still rare. Incomes are down, and unemployment is up. Some entrepreneurs and managers have achieved dramatic success, while most of their compatriots have steadily lost ground to hyperinflation. The upshot is a distended income distribution and unprecedented income inequality.

Distinctive features of late Soviet-era stratification persist. Returns to education have diminished and professionals remain relatively disadvantaged, particularly in the private sector of the economy. Stratification by place of residence and by economic branch have increased. The gender gap in earnings remains greater in the private sector than in the public sector. The emerging pattern of stratification is consistent with the specific path taken by the Russian market transition: trade, consumer services, and short-term profit making have been emphasized at the expense of production, public services, and long-term investment—a dynamic corresponding to the image of “merchant capitalism” (Burawoy and Krotov 1992). The impact of market transition on stratification is path dependent and cannot be deduced from market institutions: country-specific trajectories previous to and during the transition are decisive (Walder 1996). Market transition theory does not help us anticipate or explain these developments.

HISTORICAL BACKGROUND: ECONOMIC REFORMS IN RUSSIA

In the latter years of the Soviet era, Mikhail Gorbachev tried to reform the Soviet system from within. His economic reforms were half-hearted,

inconsistent, and poorly implemented (Goldman 1994; Aslund 1995, chap. 2). Perhaps the most significant steps were the May 1987 Law on Individual Labor Activity, which paved the way for individual and cooperative private enterprise, and the May 1988 Law on Cooperatives, which removed many restrictions on cooperatives contained in the first law. These measures quickly led to a burgeoning cooperative sector and a rise in individual contract work (Jones and Moskoﬀ 1991). But regulations still prohibited coops from hiring labor or owning the productive assets they used. Additional legislation removed these restrictions in Russia during 1990, and the right to private property in productive assets was established by 1991 (Aslund 1995, pp. 223–26).

The Soviet Union collapsed soon after the failed August 1991 coup. Yeltsin ended Communist Party control of the state and economy. The Soviet economic system had been built upon central administration, state ownership, and cradle-to-grave security. In January 1992, the Yeltsin government introduced a package of “shock therapy” reforms designed to abruptly replace the state socialist economy with market institutions and practices. Since that time, many rudiments of a market economic system—including free prices and trade, legalized ownership of productive assets, enterprise autonomy, and bankruptcy—have replaced pervasive controls, restrictions, and boundless state subsidies. The Russian economy of 1996 differs dramatically from the Russian economy of 1991, even if reforms have been fitfully implemented and some features of the Soviet economy persist.

The January 1992 reforms freed 90% of retail prices overnight; most remaining prices followed in the ensuing months. Directives and legislation rapidly liberalized trade and exchange rates, cut the state budget (reducing defense expenditures by 68%), curtailed subsidies, and established provisions for enterprise bankruptcy. The shock of price liberalization—induced inflation. The Central Bank of Russia’s loose monetary policy spurred it on to hyperinflation. A wage-price spiral ensued. Prices increased 1,100% during 1992. Wages kept pace with that inflation according to official statistics on mean monthly earnings (“Sotsial’no-ekonomicheskoe polozhenie rossiiskoi federatsii v 1992 goda” 1993), although our data cast doubt on the official numbers. Cash shortages and the diminished real value of state subsidies led to the so-called “arrears crisis” of the summer of 1992, when some enterprises could not pay for goods delivered to them and suppliers could not meet their payrolls. Many enterprises found they could not pass the rising costs they were facing down the production chain. They had no alternative but to curtail output. The economy contracted. GDP declined by 18% in 1992, 12% in 1993, 15% in 1994, and 4% in 1995 (“Sotsial’no-ekonomicheskoe polozhenie

rossiiskoi federatsii v 1992 godu" 1993; "Sotsial'no-ekonomicheskoe polozhenie rossiiskoi federatsii v ianvare-iuli 1993 goda" 1993; "Sotsial'no-ekonomicheskoe polozhenie rossii v pervom polugodii 1994 goda" 1994; "Sotsial'no-ekonomicheskoe polozhenie rossii v pervom polugodii 1995 goda" 1996). By January 1996, the Russian economy was only three-fifths as large as it had been just four years earlier. Investments declined by almost half in 1992, by 12% in 1993, 23% in 1994, and 13% in 1995. Unemployment rose from almost zero at the end of 1991 to approximately 7% of the labor force by October 1995 (Aslund 1995, p. 278; "Sotsial'no-ekonomicheskoe razvitie rossii v 1995 godu" 1996, p. 76). Involuntary furloughs and delays in payments of wages and salaries also became endemic and remain so today (Perova 1996).

Macroeconomic stabilization still eludes Russia's grasp. Inflation remains at a rate that few Western states would tolerate (though below its 1992 rate). Liberalization remains official policy, but some government intervention has been unavoidable. The federal government grants emergency credits to large enterprises on the brink of failure. It attempted to control wages via "excess wage" taxes. Some local governments intervened to control prices. But the basic principles of liberalization remain intact. The state now considers intervention in the economy only when emergency circumstances warrant.

Some indicators are actually positive. The supply of consumer goods has increased, the private sector is expanding rapidly, and personal services are now available (Aslund 1995; OECD 1995). The military consumes fewer resources.

Privatization complements stiff monetary policy. The plan to transfer ownership and governance of economic enterprises from the state into private hands was designed in 1992 and mostly implemented by June 1994. By that time over 75% of small-scale enterprises were privatized through direct competitive bidding or lease buyouts. Another 49,000 medium- and large-scale enterprises, forming 60% of industrial assets, had completed or were undergoing "mass" privatization (OECD 1995, p. 77).¹

Control changed less than ownership did. In most cases, managers gained control over the enterprises they ran during the Soviet era. Some

¹ Mass privatization consisted of four phases: corporatization, closed subscription, voucher privatization, and cash privatization. Enterprises could choose their own mix of closed subscription, voucher purchases, and outside cash purchases. For more detailed description on the design and implementation of the privatization plan, see Rutland (1994), Boycko, Shleifer, and Vishny (1995), Blasi, Kroumova, and Kruse (1997), and OECD (1995, chap. 4).

management teams acquired full ownership. More often, where the work collective as a whole obtained the majority interest, managers retained *de facto* control (Webster 1994; Boycko et al. 1995, chap. 5; OECD 1995, chap. 4). Managerial control has led to accusations of corruption and raised the potential for asset-stripping (Intriligator 1994; Rutland 1994). While these scandals have produced headlines, inertia has been more of a problem on the ground. Managers in privatized firms resist market-oriented behaviors (Kharkhordin and Gerber 1994; Webster 1994; Blasi et al. 1997). Few consider dismissing workers to be a legitimate way of cutting costs; most accede to wage and other demands from workers. Their response to economic pressures is to bargain with suppliers and to secure payment from customers (Boycko et al. 1995, chap. 5).

Privatization has changed ownership. By the end of 1994, 65% of officially registered state and municipal enterprises were transferred to private ownership (Aslund 1995, p. 273). Surprisingly few new private enterprises opened for business—probably due to barriers posed by criminals and corrupt officials (Aslund 1995, p. 263). Nonetheless, the rapid creation of a large private sector represents a significant transformation of the Soviet economy.

MARKET TRANSITION THEORY AND THE NEW RUSSIA

Under the Soviet system, central authorities prescribed wage ranges for all ranks of all occupations. Wage policies were decidedly egalitarian—especially from the late 1960s onward—in contrast to prevailing patterns in market economies (Connor 1979; Flakierski 1993). Such inequalities as there were tended to be based on economic branch or regional differences (Zaslavsky 1982). Employees of all ranks in defense and heavy industry, for example, received higher wages than their counterparts in light industry or services. Managers enjoyed higher wages than other categories of workers according to localized Soviet surveys of the 1970s (Yanowitch 1977; Connor 1979). Human capital mattered little; skill differentials played a role in earlier periods, but by the late 1980s the average wages of manual and nonmanual employees were barely distinguishable (Connor 1991). Gorbachev used the absence of rewards for skills as a justification for *perestroika*.

Education accounted for only small differences in earnings: in the Russian Federation in 1989, those with a university (VUZ—the Russian acronym for “higher educational institution”) degree earned on average only 17% more than those with a secondary school diploma and only 42% more than those with less (Flakierski 1993, p. 39). Soviet women on average earned only slightly more than 70% of men’s wages in 1984, 1989, and

1991 (Flakierski 1993, p. 34; Gradolph 1994). In sum, the Soviet system of wage distribution was characterized by unusually low returns to education and human capital and unusually high returns to privileged economic branches and regions. Managerial compensation and gender differentials were comparable to those in the West.

Nee's market transition theory posits two principal outcomes of market reform: rewards to factors that enhance productivity, especially entrepreneurial initiative and education, are enhanced, while rewards to political position (redistributive power) are reduced (Nee 1989, 1991, 1996; Nee and Matthews 1996). Together, these offsetting developments should result in little change in the overall level of economic inequality, at least initially.

Much of the debate over market transition theory concentrates on the proposition regarding returns to redistributive power, operationalized as the hypothesis that the income advantages of cadres diminish as market penetration increases (Nee 1991, 1996; Róna-Tas 1994; Bian and Logan 1996; Parish and Michelson 1996; Walder 1996). This issue is less germane to post-Soviet Russia and the other "capitalist oriented economies" in Eastern Europe (Szelenyi and Kostello 1996). After the collapse of the administrative economic system, former cadres may well have special access to entrepreneurial opportunities due to their embeddedness in social networks (Róna-Tas 1994) or their technocratic qualifications (Szelenyi and Kostello 1996). But these advantages cannot be based on power positions within the redistributive system, since the redistributive system as such no longer exists. Thus, Nee's "declining significance of redistributive power" thesis might apply to developments during the Gorbachev era when the administrative system remained intact but is irrelevant to developments in Russia since January 1992.⁴

HYPOTHESIS 1.—*Market transition increases returns to education.*

Nee's thesis regarding returns to education does apply to post-Soviet Russia. Nee reasons that in contrast to administrative systems, markets directly reward individual productivity by permitting producers to bargain over the price of their labor or product: "This is likely to be reflected in higher returns of education, which is among the best indicators of human productivity" (Nee 1989, p. 666). This hypothesis implies three predictions regarding the returns to education in post-Soviet Russia:

⁴ As Nee correctly notes, the finding that in Hungary *former* cadres were more likely to become entrepreneurs and more successful as entrepreneurs (Róna-Tas 1994) does not refute market transition theory, which claims only that cadres as such (i.e., not as entrepreneurs) lose out in the transition to a market-based system (Nee 1996, p. 915, n.3; see also Szelenyi and Kostello 1996).

HYPOTHESIS 1A.—*The total effect of education on earnings in Russia increases from 1991 to 1995.*

HYPOTHESIS 1B.—*At any given time, the returns to education in the private sector of the economy exceed the returns to education in the state sector.*

HYPOTHESIS 1C.—*The returns to education in the private sector in 1995 exceed the returns to education in the private sector in 1991.*

"Market transition theory turns on the extent to which markets replace hierarchies in the allocation of resources" (Nee 1991, p. 268). Returns to education should increase as product and labor markets become more firmly entrenched and privatization advances. Have they? The breakup of the Soviet state removed the backbone of the administrative system, but a fully functioning market economy did not immediately arise to take its place. While exchange of products and labor became formally free, structural barriers remained. Moreover, privatization was implemented only gradually. So one of the initial conditions of market transition theory is not completely met in Russia. That noted, it is also important to recognize that Russia moved away from hierarchy and toward markets throughout the first five years of the post-Soviet era.

Like private firms, state enterprises are forced to operate in increasingly marketlike conditions with the collapse of the administrative system. However, the fortunes of state enterprises may still depend on processes of political bargaining for state credits and subsidies rather than on profits and losses. Therefore, there is less incentive to link individual productivity to wages. The returns to education should be greater in private firms.

As discussed above, the privatization of an enterprise may not lead immediately to restructuring. But with time it may be expected that management in private enterprises must reorient decisions, including decisions regarding pay schedules, toward market criteria (Murrell 1996). If the time elapsed since privatization of state firms is positively associated with increased returns to education, the overall returns to education within the private sector should increase with time.

HYPOTHESIS 2.—*Market transition decreases differences between men and women.*

Market transition theory follows the usual claims that markets reward individual productivity. From this general proposition (and their observations in China) Nee and Matthews (1996, pp. 428–29) derive the proposition that the gap between men's and women's wages will narrow "the more extensive the shift to markets and the higher the rate of economic growth." If the two conditions are additive, that is, if market transition can affect the gender gap in wages in the absence of economic growth, then market transition theory implies a falling gender gap in Russia, too.

It also implies that returns to productive assets, especially education, will become similar for men and women.

HYPOTHESIS 2A.—The total effect of gender on earnings in Russia is less in 1995 than it was in 1991.

HYPOTHESIS 2B.—At any given time, Russian women's earnings are closer to Russian men's in the private than in the state sector.

HYPOTHESIS 2C.—By 1995 women's returns to education should approximate those of men in the private sector.

HYPOTHESIS 3.—Proprietors benefit from market transition.

The more widespread and firmly embedded market institutions become, the more secure and advantageous the opportunities they provide entrepreneurs. This leads to the following predictions.

HYPOTHESIS 3A.—Proprietors' absolute incomes rise from 1991 to 1995.

HYPOTHESIS 3B.—Proprietors' incomes rise in comparison to employees' wages from 1991 to 1995.

HYPOTHESIS 4.—Managers benefit from market transition.

Curiously, all sides in the market transition debate have ignored managers as a distinct group, despite their obvious importance in the political economy both pre- and posttransition. The silence may stem from the ambiguity of managers' dual position in state socialism, which combined elements of both "cadre" and "entrepreneur."⁵ While acknowledging the complexity of the issue—to which we cannot do justice here—we derive two predictions that parallel those for proprietors for heuristic purposes.

HYPOTHESIS 4A.—The absolute earnings of managers in the private sector rise from 1991 to 1995.

HYPOTHESIS 4B.—Private-sector managers' earnings rise in comparison to the earnings of public-sector managers from 1991 to 1995.

Our reading of market transition theory predicts a shakeout of managers who owed their position and high earnings mainly to their political capital (cadre status) under the old system. The dissolution of the redistributive apparatus should remove the principal positional basis for the high earnings of such managers. On the other hand, many managers de-

⁵ Their duties included administering, coordinating, and monitoring activities within their firms. Yet they were joined at the hip to the redistributive bureaucracy: beholden to it for their positions—not to mention suppliers, distributors, and investment capital—they were also advocates within it for their organizations (Burawoy and Lukacs 1992). The dissolution of the state ministries comprising the Soviet redistributive apparatus freed state sector managers from both the burdens and the blessings associated with this tie. The balance of burden and blessing most likely varied considerably across managers, depending on their individual capacities, their positions within their firm, and their firm's characteristics.

veloped proto-entrepreneurial skills in the Soviet system—bargaining, brokering, administering operations, and so forth. Market transition should enhance the rewards accruing to managers who deploy these skills—not to mention their education and social capital (contacts)—to raise their firm's productivity. Thus, some managers should experience decreased earnings and others increased earnings, with the likely net result being little aggregate change in managerial earnings in the state sector. However, in the private sector, market transition theory predicts a more rapid incarnation of a managerial labor market, leading to faster bidding up of managerial salaries than in the state sector.

HYPOTHESIS 5: Professionals benefit from market transition.

As in China (Walder 1995a), professionals were the most disadvantaged occupational group in the Soviet Union relative to their Western counterparts. For this reason, many thought they were the social group most likely to oppose the Soviet system (Parkin 1972), and by most accounts they provided the largest proreform constituency during the Gorbachev era (see Gerber 1995, chap. 2). Markets should reward their expertise. Professionals' gains in income should exceed those of other highly educated persons who find employment that is not as closely associated with their credentials.

HYPOTHESIS 5A.—The gap between professionals' earnings and those of unskilled workers increases from 1991 to 1995.

HYPOTHESIS 5B.—By 1995, professionals' earnings advantage over unskilled workers in the private sector exceeds that in the state sector.

HYPOTHESIS 5C.—Professionals' earnings increase faster in the private sector than in the state sector.

HYPOTHESIS 6.—Market transition increases returns to manual skill.

Manual skill is another form of human capital that enhances productivity. Therefore, the same logic should apply to manual skill that applies to returns to education: more skill should lead to higher earnings.

HYPOTHESIS 6A.—The gap between the earnings of skilled manual workers and unskilled workers increases from 1991 to 1995.

HYPOTHESIS 6B.—Skilled workers' earnings advantage in the private sector exceeds their earnings advantage in the state sector.

HYPOTHESIS 6C.—Skilled workers' earnings advantage in the private sector increases from 1991 to 1995.

An important caveat to this kind of application of Nee's theory is its reliance on inference instead of observation. The theory is powerful because it predicts individual-level patterns. But without direct observation of firms in context, we can neither be sure that marketization is fully in force in Russia nor that the proposed mechanisms are the actual causes (Guthrie 1997).

INCOME INEQUALITY IN RUSSIA: PREVIOUS RESEARCH

Highly aggregated official Russian data show falling real wages (as nominal pay hikes fail to keep pace with inflation), increasing proportions of the population living below the official poverty line, and surging income inequality (see, e.g., Samodorov 1992; "Sotsial'no-ekonomicheskoe polozhenie rossii v pervom polugodii 1994 goda" 1994, p. 62; Osipov, Levashev, and Lokosov 1995, pp. 32–33). But these aggregate data hint at more than reveal how wealth and poverty are allocated.

The few existing analyses of post-Soviet stratification have produced inconclusive or contradictory results. One study finds that "intellectuals, managers of various ranks, highly skilled specialists . . . businessmen, entrepreneurs, or, more broadly, the self-employed population" enjoy the highest earnings in the new environment, while unskilled workers, pensioners, the unemployed, agricultural workers, and, lately, skilled workers face drastically lower incomes (Khakhulina and Tucek 1996, pp. 28–32). Yet others claim that many groups of professionals, particularly scientific-technical workers and academics, are among the biggest losers (Anurin 1995) or that the "mass intelligentsia" share the poverty of unskilled agricultural and manual workers (Zaslavskaja 1995). Yet elsewhere in the same article, Zaslavskaja (1995, p. 34) states: "A typical representative of the affluent stratum of Russian society is a male, 30–40 years old, who has completed a higher education, has high professional qualifications, and is involved in intense professional activity." Most studies are merely descriptive. Researchers have described the composition of the poorest (Samodorov 1992; Mroz and Popkin 1995) and wealthiest (Kryshtanovskaja 1995) groups.

The absence of multivariate analysis frustrates any attempt to draw inferences. We cannot tell whether Zaslavskaja's "typical representative of the affluent stratum" owes his affluence to his maleness, his youth, his education, his occupation, or his employment in the private sector because her summary refers to a series of bivariate comparisons. Furthermore, these studies tend to focus on data gathered at one moment in time.⁶ In the volatile setting of the transition economy, multiple data points must be considered to distinguish steady patterns from transient developments. Moreover, to assess whether reforms have altered employment and income patterns, we must compare post-Soviet results with a prereform baseline.

⁶ An exception is Mroz and Popkin (1995), but their focus is explicitly on poverty.

DATA AND METHODS

We analyze data from five surveys, conducted in February 1991, February 1992, March 1993, July 1994, and January 1996. These data permit us to track developments in employment patterns, income levels, and income differentials over the reform period from January 1991 to December 1995 and to evaluate the hypotheses implied by market transition theory. The 1991 data were gathered as part of a collaborative effort of (then) Soviet and American researchers in connection with the Comparative Class Structure and Consciousness Project (CCSCP). The Institute of Sociology in Moscow supervised data collection.⁷ The 1992 data were collected by the Moscow-based All-Russian Center for the Study of Public Opinion, widely known by its Russian acronym, VTsIOM, as part of the 1992 International Social Survey Programme (ISSP). The remaining data sets were also collected by VTsIOM as part of a continuing series of monthly "omnibus" economic surveys initiated in March 1993.⁸

Each survey drew a multistage stratified cluster sample of a target population. The 1991 CCSCP study limited the target population to European Russia. The remaining surveys sampled most of the Russian Federation; they excluded only war zones and the most remote settlements.⁹ Some of the contrasts between the Soviet and post-Soviet periods found in our data may be artifacts of differences between the European Russian population and the population of the entire Russian Federation. However, we doubt that these differences are great enough to distort our comparisons,

⁷ For details on sampling and data collection for the CCSCP survey in Russia, see Hout, Wright, and Sánchez-Jankowski (1992).

⁸ VTsIOM has been conducting large-*N* surveys using modern sampling techniques since 1989 and has earned a reputation as the preeminent survey research operation in Russia. Its data are widely cited by scholars (e.g., Ashund 1995) and international organizations (e.g., OECD 1995) as authoritative, and VTsIOM proved singularly accurate in forecasting the outcome of both rounds of the recent Russian presidential elections. The VTsIOM data we have at hand are the most reliable available.

⁹ The Northern Caucasus was excluded due to political and military conditions there, and some northern territories were excluded due to inaccessibility. The sampling procedure for the VTsIOM surveys is as follows: Apart from the excluded areas, all settlements—defined as either "urban population points" or "rural administrative areas"—are divided into approximately 65 strata according to region, a proxy for ethnic composition, size, and administrative status. Target sample volume is then distributed across strata proportionately to size of adult population (based on official statistics from the 1989 census). The number of primary sampling units (PSUs) required within each stratum to attain its target sample size without exceeding 45 respondents per PSU is determined. PSUs are selected from within strata using systematic selection with probabilities proportionate to size. Secondary sampling units are electoral districts (in urban PSUs) or "villages" (in rural PSUs), chosen using systematic selection. Dwellings are sampled randomly within electoral districts using a random walk algorithm and respondents within dwellings using the Kish method.

and we address the major source of potential distortion by weighting the data sets by demographic variables. The VTsIOM data sets came with weights to correct for discrepancies between sample distributions and the most reliable official statistics on education, gender, age, and type of residence. We devised similar weights for the class project sample. We use the weights in all of our calculations reported here so that we can rule out random fluctuations in demographic composition across samples as the source of the changes we observe.

The variables in our analysis include respondents' demographic characteristics (education, gender, age, type of residence), employment variables (current employment status, employment in the state or private sector, branch of the economy in which currently employed), occupational class (based on primary occupation), and earnings from primary occupation in the month preceding the survey. Because the surveys were not designed to be comparable, we had to standardize several of our measures across disparate coding schemes (esp. occupation, economic branch, and region; see appendix).

Demographic variables exhibit only minor variations across samples, as shown in table 1. This reflects the influence of our weights. Trends in workforce status capture the growth of unemployment and nonemployment following the onset of market reforms (Mikhalev 1996).¹⁰ Our data document the steady growth of the private sector during the course of the reform period, as well as the growth of services, the most dynamic branch in the new Russian economy (OECD 1995, p. 5).

The educational summary measure "years of schooling completed" fails to capture the complex credentialing and vocational preparation that constitutes secondary and higher education in Russia. We use a set of discrete categories to capture the most important types of educational outcome in the Soviet educational system (see Gerber and Hout 1995): a VUZ degree, complete secondary (combining "general" secondary and "special" secondary, the latter having a more technical component), vocational (PTU and FZO, both designed to impart manual skills), and less than secondary (none of the above).

To classify occupations, we modified the widely used Erikson-Goldthorpe (1992) schema to our purposes. Because of our interest in the role of managers in Soviet and post-Soviet Russia, we separate managers into their own class (we had to combine upper and lower managers due to

¹⁰ Women remained out of the labor force in greater proportion than men. In February 1991, 18.4% of male respondents were pensioners or were otherwise not in the labor force (i.e., keeping house, attending school, unable to work, or "other"), growing to 24.7% in January 1996. The corresponding figures for women respondents are 31.5% and 41.2%.

TABLE 1

DESCRIPTIVE STATISTICS, RESPONDENTS 18 AND OLDER: RUSSIA, 1991-96

Variable	February 1991	February 1992	March 1993	July 1994	January 1996
Education:					
<i>N</i>	2,105	1,865	3,879	2,884	2,340
Less than secondary	36.3	36.3	28.7	29.5	30.8
Vocational (PTU and FZO)	9.6	13.9	16.4	17.5	16.3
Complete secondary (including special secondary)	42.0	37.7	41.9	41.3	40.7
VUZ degree	12.1	12.1	12.9	11.7	12.1
Sex:					
<i>N</i>	2,105	1,865	3,879	2,884	2,340
Men	43.9	44.8	45.2	45.4	45.5
Women	56.1	55.2	54.8	54.6	54.5
Age:					
<i>N</i>	2,105	1,865	3,879	2,884	2,340
18-25	14.4	14.4	14.7	16.5	15.5
26-40	32.4	34.0	33.7	34.1	34.6
41-55	26.3	25.7	23.3	22.8	22.5
Over 55	26.8	25.9	28.3	26.6	27.4
Residence type:					
<i>N</i>	2,105	1,864	3,878	2,884	2,340
Moscow and St. Petersburg	9.6	8.3	8.9	9.2	9.3
Other large cities (> 500,000)	20.1	19.8	19.9	17.4	18.8
Medium and small cities (100,000-500,000)	19.1	28.4	21.9	17.9	18.4
Small towns (< 100,000)	25.1	16.6	22.8	29.9	27.9
Rural towns and villages	26.1	26.9	26.4	25.7	25.6
Workforce status:					
<i>N</i>	2,104	1,275	3,879	2,884	2,426
Employed	67.0	70.3	62.0	59.2	55.1
Unemployed	0.7	1.1	4.1	7.3	8.2
Out of labor force	13.2	8.5	9.9	8.9	12.0
Pensioner	19.0	20.1	24.0	24.7	24.6
Sector:					
<i>N</i>	1,389	1,249	2,378	1,683	1,307
State	96.2	92.2	82.4	81.1	79.6
Private	3.8	7.8	17.6	18.9	20.4
Economic branch:					
<i>N</i>	1,412	1,234	2,388	1,694	1,310
Industry*	36.7	47.6	44.1	41.1	41.6
Agriculture†	12.7	15.0	14.2	12.8	13.6
Services‡	10.4	12.8	15.6	18.3	15.6
Cognitive§	22	18.7	18.1	19.4	18.7
State/Financial 	7.5	4.1	6.2	7.2	8.6
Other	10.7	2.0	1.8	1.1	1.8

TABLE 1 (Continued)

Variable	February 1991	February 1992	March 1993	July 1994	January 1996
Classes:					
<i>N</i>	1,390	1,196	2,365	1,649	1,260
Ia/IIa. managers	7.1	4.4	7.1	4.8	3.2
Ib/IIb. professionals	17.0	19.2	16.1	17.5	15.8
IIIa. upper routine nonmanual	5.0	3.9	5.3	7.3	8.1
IIIb. lower routine nonmanual	18.4	14.1	17.7	16.2	18.8
IV. proprietors	1.7	1.3	1.6	1.4	1.4
V. technicians and foremen	8.3	7.6	8.3	8.3	8.8
VI. skilled manual	17.7	27.6	25.4	26.9	25.3
VIIa. semi- and unskilled, industry	19.2	20.1	17.9	16.1	16.9
VIIb. semi- and unskilled, agriculture	5.5	1.8	0.7	1.5	1.7

NOTE.—Data are weighted by demographic variables. Weights are calculated so that valid *N* is equal to the number of unweighted observations. All numbers except *N*s are percentages.

* Includes transport, construction, communications.

† Includes other primary.

‡ Includes domestic, consumer, trade, insurance.

§ Includes education, health, science, culture.

|| Includes state organs, social organizations, finance.

lack of data on size of enterprise) and treat *all* proprietors as a separate class. After preliminary analyses showed no income differences between the upper and lower professional occupations, we eliminated the distinction.

The class distributions are fairly stable across the five data sets. The 1991 data contain a larger proportion of managers and agricultural workers and a smaller proportion of skilled manual workers than most of the remaining samples. Reforms may well have depleted managerial ranks, although coding inconsistencies (see appendix) and the 1991 focus on European Russia may also be factors.

The most revealing occupational data in table 1 are the numbers for proprietors. The self-employed remain a trivial segment of the Russian labor force—less than 2% of the currently employed sample—throughout the early 1990s. Apparently, privatization has created far more collective and institutional owners in the Russian economy than individual owners. New private enterprises remain a rather insignificant economic element. Aslund (1995) corroborates the sluggish development of new private enterprises in Russia, particularly striking in comparison to Eastern Europe. He attributes the slow pace to obstacles imposed by criminals and corrupt officials (Aslund 1995, p. 263). That might be part of it, but legal obstacles including license requirements and tight credit should not be ruled out (Burawoy and Krotov 1992).

Our hypotheses suggest that the effects of education, class, gender, residence type, and economic branch on earnings covary with the accomplished level of market transition. We operationalize the "accomplished level of market transition" as the amount of time elapsed since February 1991, measured in years. This measure is less than perfect because it imposes linearity on a multidimensional and somewhat sporadic process. Nonetheless, approximating the growth of market institutions in Russia from early 1991 through the end of 1995 by using a simple straight line is intuitively satisfactory, particularly given the timing of the subsequent data points. January 1992 saw the introduction of the Russian government's stabilization program, in effect compressing an entire year's worth of reforms into the month preceding the collection of our second data set (in February 1992). Then the privatization program was initiated and began to show results by March 1993. Privatization continued apace through our July 1994 data point and through January 1996. Although the pace of privatization decelerated during 1994 and 1995 (see "Sotsial'no-ekonomicheskoe razvitie rossii v 1995 godu" 1996, p. 80), by this time previously privatized enterprises would be getting market cues regarding how well they were doing.¹¹

The composition of each class by gender, education level, and economic branch changed little in five years (table 2). Professions, the upper and lower routine nonmanual occupations, and the ranks of technicians and foremen have a higher proportion female than does the labor force as a whole. Men are overrepresented among managers, proprietors, and skilled and unskilled manual workers. This pattern of gender segregation is familiar (Charles 1992). University degrees are not as widespread in Russia as they are in the United States and some other developed countries; compare the 15% and 16% estimates in our data with the 30% in the United States (Mare 1995). The professions and management have many VUZ-degree holders. Very few service and blue-collar workers have VUZ degrees. Professionals are highly concentrated in the "cognitive" branches of the economy (education, health, science, culture). The few proprietors are concentrated in the "service" branches (trade, services, and insurance). These activities require little capital compared to the prerequisites of starting up a manufacturing enterprise. Consequently, they have pro-

¹¹ More complex, composite measures of "degree of marketization" might be developed based on standardized measures of different aspects of marketization (percentage employed in the private sector, proportion of enterprises in private hands, relative contribution to GDP of state and private sectors, etc.) These measures may be less precise than they at first appear, based as they would be on unreliable official data. We doubt that they would seriously depart from temporal linearity across our data points; certainly they would show a monotonic increase in level of marketization over the time period covered by our data.

TABLE 2
CHARACTERISTICS OF CLASSES (%): RUSSIA, 1991-96

Characteristic	February 1991	February 1992	March 1993	July 1994	January 1996
Proportion women	51	50	48	50	49
Ia/IIa. managers	41	45	37	29	38
Ib/IIb. professionals	62	72	66	69	69
IIIa. upper routine nonmanual	71	38	47	50	43
IIIb. lower routine nonmanual	90	88	84	81	80
IV. proprietors	25	3	10	8	20
V. technicians and foremen	68	66	65	79	86
VI. skilled manual	23	34	33	34	26
VIIa. semi- and unskilled, industry	24	23	20	23	20
VIIb. semi- and unskilled, agriculture	55	40	0	27	22
Proportion with VUZ degree	15	15	16	15	16
Ia/IIa. managers	40	32	50	43	61
Ib/IIb. professionals	50	48	53	52	59
IIIa. upper routine nonmanual	24	26	11	14	14
IIIb. lower routine nonmanual	2	5	4	4	3
IV. proprietors	26	37	38	30	26
V. technicians and foremen	17	13	15	11	11
VI. skilled manual	1	4	2	3	4
VIIa. semi- and unskilled, industry	0	3	1	1	1
VIIb. semi- and unskilled, agriculture	1	0	0	3	0
Proportion in cognitive branches	22	19	18	20	19
Ia/IIa. managers	23	8	18	6	16
Ib/IIb. professionals	58	45	49	50	48
IIIa. upper routine nonmanual	5	12	10	8	7
IIIb. lower routine nonmanual	23	32	22	15	20
IV. proprietors	0	5	1	5	1
V. technicians and foremen	52	48	46	61	58
VI. skilled manual	3	2	2	6	1
VIIa. semi- and unskilled, industry	5	3	3	6	5
VIIb. semi- and unskilled, agriculture	9	0	0	0	0
Proportion in service branches	11	13	16	18	15
Ia/IIa. managers	7	20	15	31	24
Ib/IIb. professionals	1	6	4	6	6
IIIa. upper routine nonmanual	16	13	21	34	21
IIIb. lower routine nonmanual	24	33	36	39	37
IV. proprietors	7	47	65	69	68
V. technicians and foremen	3	5	7	4	1
VI. skilled manual	9	11	11	10	8
VIIa. semi- and unskilled, industry	14	7	10	17	11
VIIb. semi- and unskilled, agriculture	1	0	0	0	0

NOTE —Currently employed, weighted sample.

duced more visible results than manufacturing has during the first five years of reform.

Our analysis consists of two steps. First, we present and discuss trends in employment and earnings found in our data. In addition to the aggregate trends, we examine the group-specific marginals to see how different groups have fared in the transition process. The trends in our data provide an important empirical baseline, since descriptive results in this form have not been presented elsewhere. We then turn to multivariate regression analyses in order to test our hypotheses regarding the determinants of earnings more directly.

TRENDS IN EMPLOYMENT

Since the emergence of co-ops in the late Gorbachev era, the private sector has promised the greatest opportunities for Russians to improve their economic position. Our data show that independence from the state sector has worked to some peoples' advantage. At the same time, the appearance of significant unemployment means that other Russians have lost out in the transition process. The old system insisted on full employment and the new system has yet to provide realistic unemployment benefits (Mikhalev 1996).

Private Sector Employment

Proprietors all work in the private sector by definition (table 3). But they are a surprisingly small fraction of the labor force (less than 2%). Managers and routine nonmanual employees moved into private sector employment most rapidly (or held their jobs at enterprises that became privately held). Though managers' private sector employment rate did not grow quite as rapidly as the overall rate, it stood 50% higher than the overall rate by January 1996. Professionals, in contrast, still tend disproportionately to work in the shrinking state sector, as do technicians and unskilled workers. Skilled workers remain close to the overall rate of private sector employment in each period. These patterns suggest that ownership change is a more important source of change than individual mobility from state-run firms to privately held ones.

We have some evidence of private firm formation. Younger cohorts of workers have been more likely than their elders to work in the private sector, and the gap between younger and older workers increased over time. The two "capital" cities (Moscow and St. Petersburg) have consistently offered more private sector opportunities than other places, and private sector employment has remained lowest in rural areas. Finally, private sector growth has been most rapid by far in the service branch:

TABLE 3
PRIVATE SECTOR EMPLOYMENT (%): RUSSIA, 1991-96

Variable	February 1991	February 1992	March 1993	July 1994	January 1996
Total private sector employment	4	8	18	19	20
Classes					
Ia/IIa. managers	11	20	27	33	31
Ib/IIb. professionals	1	5	10	11	13
IIIa. upper routine nonmanual	6	27	17	38	23
IIIb. lower routine nonmanual	1	4	16	22	27
IV proprietors	100	100	100	100	100
V. technicians and foremen	0	2	7	9	6
VI skilled manual	1	6	18	19	23
VIIa. semi- and unskilled, industry	1	5	18	11	14
VIIb. semi- and unskilled, agriculture	3	0	22	11	7
Education:					
Less than secondary	1	6	21	15	16
Vocational (PTU and FZO)	3	12	16	15	27
Complete secondary (including special secondary)	5	6	16	20	24
VUZ degree	7	11	18	25	23
Sex:					
Men	6	11	21	22	23
Women	2	4	14	15	18
Age:					
18-25	4	10	24	28	26
26-40	4	10	20	22	25
41-55	3	5	14	12	16
Over 55	3	2	11	9	3
Residence type:					
Moscow and St. Petersburg	5	8	26	36	37
Other large cities (> 500,000)	2	16	18	20	22
Medium and small cities (100,000-500,000)	4	6	18	22	21
Small towns (< 100,000)	6	10	12	17	19
Rural towns and villages	2	3	19	11	14
Economic branch:					
Industry	2	8	13	14	19
Agriculture	2	2	23	10	9
Services	1	18	42	55	62
Cognitive	1	3	5	5	5
State and financial	6	10	6	5	1
Other	21	23	34	47	32

TABLE 4
UNEMPLOYMENT (%). RUSSIA, 1991-96

	February 1991	February 1992	March 1993	July 1994	January 1996
Weighted <i>N</i> (employed + unemployed)	1,427	1,281	2,563	1,916	1,529
Total proportion unemployed	1	2	6	11	13
Education:					
Less than secondary	2	1	8	14	19
Vocational (PTU and FZO)	1	3	5	11	14
Complete secondary (including special secondary)	1	1	7	11	12
VUZ degree	1	1	3	8	5
Sex:					
Men	2	2	6	12	14
Women	1	1	6	10	11
Age:					
18-25	0	4	12	21	22
26-40	1	2	7	11	12
41-55	2	1	3	6	11
Over 55	1	1	5	5	9
Residence type:					
Moscow and St. Petersburg	3	3	8	10	12
Other large cities (> 500,000)	4	1	5	7	7
Medium and small cities (100,000-500,000)	0	1	7	12	9
Small towns (< 100,000)	0	2	7	13	18
Rural towns and villages	1	1	5	11	14

by July 1994 over half the service employees were in converted and new private firms or organizations. This helps explain the increasing private sector employment of the upper routine nonmanual occupations, which include agents and representatives in various commercial and service industries. In contrast, the "cognitive branch" (education, science, health, and culture) and "state financial branch" (public/private organizations, finance) have remained under state ownership, resulting in a low level of private sector employment among professionals.

Unemployment

We use the U.S. definition of unemployment: the unemployed have no job at present and are looking for work (table 4). Unemployment afflicts the least educated Russians and youth. By the beginning of 1996, unemployment ranged from 19% among those with less than a secondary school

TABLE 5
MONTHLY EARNINGS, SELECTED MONTHS: RUSSIA

	January 1991	January 1992	February 1993	June 1994	December 1995
Official data:					
Inflation index	1 00	8.46	100.17	1,079.25	4,362.89
Mean nominal wage	296	1,438	18,672	207,500	482,000
Mean real wage (January 1991 rubles)	296	170	186	192	110
Sample data:					
Valid <i>N</i> (weighted)	1,387	1,233	2,309	1,581	1,198
Averages:					
Mean nominal	262	1,501	19,015	188,540	578,230
Mean real (nominal/inflation index)	262	177	190	175	133
Mean log real	5.42	4.93	4.81	4.77	4.57
Geometric mean	226	139	123	118	97
Median real	220	142	120	111	92
% below 150 real rubles	18	55	64	64	73
Inequality measures:					
SD real	172	184	466	387	134
Skewness	4.13	7.00	20.19	22.71	4.18
Interquartile range	130	106	120	113	103
Top decile cut (real)	420	296	299	324	252
Bottom decile cut (real)	120	59	50	46	34
Ratio top/bottom decile cuts	3.50	5.00	6.00	7.00	7.33
SD log real53	.66	.82	.82	.77

diploma to 5% for college graduates and from 22% of 18–25-year-olds to 9% of those over 55 years old. Note though that involuntary retirement probably masks some unemployment among out-of-work 55–64-year-olds; unfortunately, our data do not distinguish between voluntary and involuntary retirements.

At the outset of the reform period, women were much less likely to be unemployed than men. One possible explanation is that jobless women were quick to simply withdraw from the labor force. As incomes have fallen, staying at home has become a luxury many cannot afford, and women's unemployment has caught up with men's in recent years. Unemployment in rural areas, initially only a fraction of that in the capital cities, quickly caught up with and overtook urban unemployment by mid-1994.

FALLING EARNINGS / GROWING INEQUALITY

Official data show how inflation has eroded Russians' earnings since prices were liberalized in January 1992 (table 5). According to our inflation

index, calculated as the cumulative changes since January 1991 in the Aggregate Consumer Price Index (*Svodnyi Indeks Potrebitel'skikh Tsen*) reported by the State Statistical Committee (*Goskomstat*), prices increased by a factor of 4,363 from January 1991 through December 1995.¹² During the same period, the officially reported mean nominal wage increased by a factor of 1,628, indicating a severe decline in real wages as prices rose four times faster than the average wage. The most dramatic decline in real wages occurred immediately following the price liberalization of January 1992. Following the initial shock, real incomes declined more slowly in 1993 (the February 1993 mean represents a temporary rebound) and in fact grew somewhat during the course of 1994, only to decline once again in 1995, according to *Goskomstat* sources (*OECD 1995*; "Sotsial'no-ekonomicheskoe razvitiie rossii v 1995 godu" 1996).

Throughout our study, we report wages in January 1991 rubles (nominal wages deflated by the inflation index).¹³ Our sample income statistics show patterns similar to the official aggregate data. The 559% increase in mean nominal wages between our February 1991 and February 1992 samples masks a *decline* in mean real wages to 65% of their February 1991 levels. By February 1993 real wages had recovered somewhat, consistent with official data, but our July 1994 sample shows a renewed decline, followed by a sharper decline through December 1995. According to our data, then, the average earnings of Russians decreased by approximately 49% from January 1991 to December 1995. If anything, our data understate the overall decrease (63% according to the official data reported in table 4).¹⁴

Most Russians experienced a substantial decline in their real wages during the market transition. The decline was not limited to the initial period of price liberalization; it continued through the end of 1995. For a sense of what the loss in real income entailed for consumption, consider the

¹² The *Goskomstat* price data on which our inflation index is based are reported in the following sources: "Sotsial'no-ekonomicheskoe polozhenie rossiiskoi federatsii v 1992 godu" (1993, p. 93), "Sotsial'no-ekonomicheskoe polozhenie rossiiskoi federatsii v ianvare-liule 1993 goda" (1993, p. 61), "Sotsial'no-ekonomicheskoe polozhenie rossii v pervom polugodii 1994 goda" (1994, p. 53), "Sotsial'no-ekonomicheskii monitoring polozhenii rossii i otidel'nykh stran mira" (1995, p. 55), and "Sotsial'no-ekonomicheskoe polozhenie rossii v pervom polugodii 1995 goda" (1996, p. 44).

¹³ For each data set, we use the month preceding data collection as the basis for computing the pertinent inflation index, since the questions on income specifically pertain to the *previous* month. Monthly income rather than annual income is the customary measure for Russians, an especially useful unit in conditions of massive inflation when annual incomes are meaningless.

¹⁴ The discrepancy may result from oversampling of higher status occupations within education groups, for which weighting did not adjust, or from the possible omission of proprietors' incomes in computations of the official mean wage (see *OECD 1995*).

burgeoning percentage below 150 rubles/month, which Mroz and Popkin (1995, p. 7) cite as a reasonable average subsistence income for one person in 1991. In 1991, 18% of Russians earned less than this standard of subsistence income; in January 1996, 73% had insufficient earnings from their job to meet it. More goods were available in 1996, but falling wages and falling subsidies kept them out of reach for most Russians.

Our data also indicate that earnings declined most steeply at the bottom end of the wage and salary distribution. From January 1991 through June 1994, the median fell by 50% compared to a 37% drop in the mean; the standard deviation increased, as did the positive skewness. Most tellingly, while the ninetieth percentile fell by 23% during this period, the earnings of the tenth percentile lost 62% of its purchasing power. Accordingly, the ratio of ninetieth to tenth percentiles—a good measure of the overall level of wage inequality—doubled.¹⁵ All these statistics point to growing inequality as some groups bore the brunt of falling real wages.

However, the most recent data point on income—from the January 1996 data set—clouds the picture somewhat. The 24% drop in the mean outstripped the 18% drop in the median, bringing the skewness down to its prereform level, and the standard deviation of the mean dipped below its prereform level. Histograms suggest that a floor effect stopped the collapse of the lowest end of real incomes (fig. 1).¹⁶ Enterprises probably cannot employ workers at wages below one-third of the poverty line (50 real rubles). Inflation continues to drag the incomes of those in the middle and upper ranges toward the floor, as indicated by higher kurtosis that replaces skew in the descriptive statistics. In June 1994, 3% of the currently employed sample earned 25 rubles or less, while 50% of the sample earned between 25 and 125 rubles. The corresponding figures for December 1996 are 4% and 62%.¹⁷

¹⁵ As a point of reference, the March 1995 Current Population Survey indicates that the ratio of the ninetieth to the tenth percentile in the United States was 5.59 in 1994 (Hout 1996)—very close to our Russian figures for all but 1994. Smeeding and Gottschalk (1996) estimate that the United States and Russia have more inequality than any other industrial nations.

¹⁶ To facilitate illustration, in fig. 1, incomes are grouped into 50-ruble categories centered around multiples of 50 and 100, and truncated at 1,000 rubles.

¹⁷ The observed decrease in inequality may also be an artifact of differential nonreporting rates. In the first three samples, over 97% of respondents answered the question regarding their income from their primary occupation. This figure dropped to 94% in July 1994 and 90% in January 1996. If nonreporters are more likely to be at the upper end of the distribution, our estimates of the means, standard deviations, and ninetieth percentiles at these points may be downwardly biased relative to the estimates for the first three data sets. To assess the possible impact of differential nonreporting on our results, we simulated the income distributions for the two most recent data sets under the assumption that excess nonreporters (those beyond the 3% who presumably do not report due to a random predisposition not to) had high in-

Trends in respondents' income from their primary occupation suggest a precipitous decline in Russians' living standards since the collapse of the Soviet state. However, many Russians have more than one job; 11% of the currently employed March 1993 respondents reported jobs in addition to their main job, rising to 15% in July 1994 and 18% in January 1996. Also, nonwage income comprises a substantial component of overall household incomes: from 21% to 23% for the currently employed respondents in the three most recent data sets. Finally, other household members may bring in income, too, a point on which our data are sketchy. Our data almost certainly overstate the decline of household income but do a much better job of monitoring the changes in personal income—the key indicator of well-being.

Economic reform has dramatically increased Russian social stratification. In 1991 it was a sleepy, low stakes game typified by the popular joke, "We pretend to work and they pretend to pay us." By 1996 the Russian stratification system was far more demanding, and the stakes were perhaps the highest among large economies (Smeeding and Gottschalk 1996). A high stakes game may or may not be fair, depending on the rules that determine the outcome. We turn to those rules for Russia by comparing the determinants of earnings across our six surveys.

DETERMINANTS OF INCOME

Russians' incomes vary by class, sector, economic branch, and demographic variables (see app. table A1). To sort out the net effects of these variables, we enter them in a (by now) standard earnings model along with selected interaction effects. We test for cross-time variation in the effects of education, class, sector, and other variables, by pooling our

comes. For the incomes of nonreporters, we substituted a value randomly drawn from $\text{nor}(y, \sigma)$, where y equals the observed ninetieth percentile income and σ equals the observed standard deviation of the eighty-fifth through ninety-fifth percentiles for the appropriate sample. We reassigned "missing" values to the appropriate proportion of these randomly selected cases so that the resulting valid percentage would resemble the 97% observed in the first three data sets. The simulation allayed our concern that the leveling effect we observed is an artifact of differential nonreporting. For the July 1994 data set, the simulated mean income (valid $N = 1,654$, 96.9%) came to 182, the simulated standard deviation to 379. For the January 1996 data set, the simulated mean (valid $N = 1,310$, 98.3%) equaled 142, the standard deviation 133. The impact of differential nonreporting was negligible within the bounds of our assumptions. More generally, underreporting may be more serious at the higher end of the income distribution across the board, but this should result in a fairly uniform downward bias on the skewness and cannot account for the clustering of respondents over time at the floor interval we observe in fig. 1.

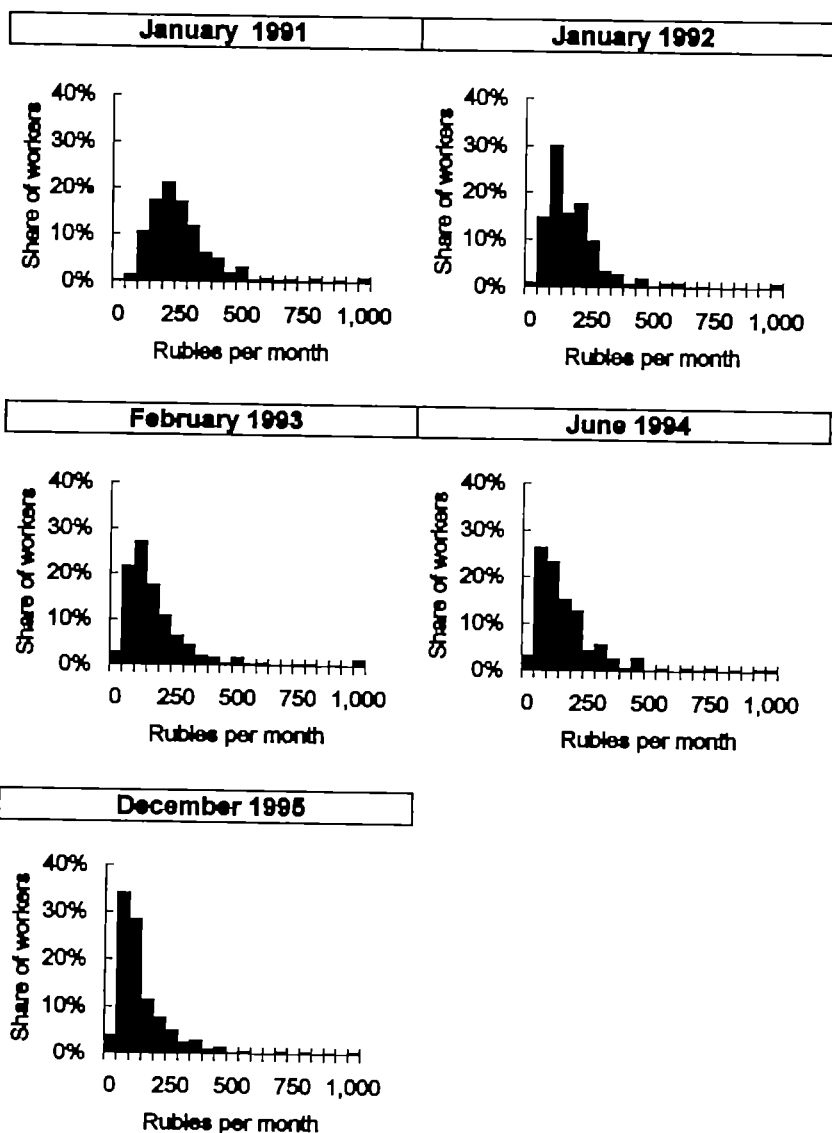


FIG. 1.—Monthly earnings of individuals in Russia (truncated at 1,000 rubles), February 1991–December 1995.

five data sets and estimating a series of OLS regression models of the form:

$$\begin{aligned}
 y_k = & \beta_0 + \sum_{t=1}^T \beta_{1t} T_{it} + \sum_{j=1}^J \sum_{k=1}^K \beta_{2jk} X_{jkt} + \sum_{j=1}^J \sum_{k=1}^K \beta_{3jk} X_{jkt} t_k \\
 & + \sum_{l=1}^L \sum_{m=1}^M \beta_{4lm} Z_{lmt} + \sum_{l=1}^L \sum_{m=1}^M \beta_{5lm} Z_{lmt} t_k \\
 & + \beta_6(\text{age}_k - 18) + \beta_7(\text{age}_k - 18)^2 \\
 & + \beta_8(\text{age}_k - 18)t_k + \beta_9(\text{age}_k - 18)^2 t_k + \epsilon_k,
 \end{aligned} \tag{1}$$

where y_k is the natural logarithm of respondent k 's earnings during the previous month (in January 1991 rubles); β_0 is the intercept; the other β s are regression coefficients; T_{it} [$t = 1, \dots, 4$] is a vector of four dummy variables capturing the main effects of period; X_{jkt} is a vector of dummy variables corresponding to J independent variables, each with $K + 1$ categories; t_k is a scalar equal to the years elapsed since January 1991;¹⁸ Z_{lmt} is a vector of dummy variables corresponding to L interactions among selected X_{jkt} , each with $K_1 + K_2 = M$ joint categories for any $j = 1, 2$; the remaining terms are continuous, polynomial specifications of the age effect, centered on 18 years of age, with both terms multiplied by t , and ϵ_k is a normally distributed error (with a constant variance σ).¹⁹

The hypotheses we derive from market transition theory correspond to predictions about the parameters in equation (1). Change in effects over time and across the state and nonstate sectors shows the impact of market transition on the Russian workforce. Our analysis evaluates these changes in terms of their combined effects on the incomes of different groups of Russians.

We estimated three sets of models (see details in app. table A2).²⁰ First,

¹⁸ The values for t are as follows: January 1991, $t = 0$; January 1992, $t = 1.00$; February 1993, $t = 2.08$; June 1994, $t = 3.33$; and December 1995, $t = 4.83$.

¹⁹ We determined that the polynomial specification of the curvilinear age effect is appropriate by examining the standardized residuals from a model including only a linear effect of age. We also used weighted least squares (WLS) to correct for minor heteroscedasticity associated with the private sector dummy variable. The WLS results closely replicate the OLS results; see the appendix for details.

²⁰ All results are based on weighted data. Poststratification weights to correct for over- and under-samples may seriously affect regression results if cases with exceptionally large weights have unusually strong influence on the regression outcomes. We were able to rule out this possibility. First, we examined the correlations (using the unweighted sample) between case weights and three measures incorporating influence saved from our model 6 regression using the weighted sample: Cook's distance, the absolute value of the studentized "deleted" residuals, and the absolute value of the standardized q -fit statistic. The correlations were small and negative, ranging from $-.025$ to $-.058$. This implies that, if anything, cases with higher weights have some-

to assess trends in the total effects of education and demographic variables, we consider a reduced form model of the effects of gender, age, residence, education, and education by gender (models 1 and 2). Then, we incorporate the mediating factor of sector, as a main effect and in interaction with education and gender (models 3 and 4). Finally, we estimate a full model including the effects of class and branch, as well as the class-by-sector interactions implied by our hypotheses 4, 5B, and 6B (models 5 and 6). In each case, we began with the complete model, including all regressor variables (models 1, 3, and 5), then we removed nonsignificant effects iteratively to arrive at a "preferred" model (models 2, 4, and 6).²¹

The numerous interaction effects among regressors, between regressors and time, and between the interactions among regressors and time make the coefficients in appendix table A2 difficult to interpret, and the logged dependent variable further complicates the picture. To simplify analysis, we convert the coefficients in models 4 and 6 into expected period-specific partial income multipliers (tables 6 and 7).²²

In table 6, the "baseline" gives the expected earnings for each period (in January 1991 rubles) of 18-year-old males living in Moscow or St. Petersburg who have a complete secondary education and are employed in the state sector.²³ The entries corresponding to each variable represent the partial income multiplier (according to model 4) associated with membership in the relevant category for the period in question. Thus, in January 1991 a woman was expected to earn 68% of a male with the same age,

what smaller influence on the weighted regression results. Seven cases had weights greater than 8.00 (ranging from 8.16 to 17.46; only 0.7% of respondents had weights over 5.00). We deleted these cases and ran model 6 again. Despite the loss of 42 weighted cases, the pattern of significance among the coefficients was unchanged.

²¹ We removed nonsignificant effects to improve the efficiency of our estimates and to facilitate computation of partial expected income multipliers. In particular, high correlations between the X_{jt} and $X_{jt}t$, and between the Z_{it} and $Z_{it}t$ inflate the corresponding standard errors. Thus, for jointly nonsignificant pairs of these variables—e.g., private \times Ib/Iib and private \times Ib/Iib \times T in model 5—we first removed the simple effect, provisionally keeping the time-interaction variable in the model. We also removed variables whose coefficients were significant according to one-tailed criterion of $P < .05$ when the sign of the coefficient did not correspond to the appropriate tail (e.g., woman \times VUZ degree in model 5). We present only the initial and final models to save space.

²² We found that the first set of reduced form models (1 and 2 in table A2) produced nearly identical results to the second set, which include the main and interaction effects of private sector employment. To avoid redundancy and save space, we bypass discussion of the first set, since more hypotheses are tested in the second set. Our observations regarding the effects of education, gender, and their interaction apply equally to both.

²³ For each period t , the baseline is obtained as $\exp(\beta_0 + \beta_{1t})$.

TABLE 6

EARNINGS MULTIPLIERS BY DEMOGRAPHIC VARIABLES AND SECTOR, REDUCED FORM

Variable	January 1991	January 1992	February 1993	June 1994	December 1995
Baseline*	206.61	142.48	139.03	156.08	151.90
Woman68	.66	.64	.61	.58
Age (40 vs. 18)†	1.54	1.41	1.29	1.17	1.03
Locality (Moscow and St. Petersburg):					
Other large cities (> 500,000)	1.00	.98	.96	.94	.91
Medium and small cities (100,000–500,000)	1.00	.96	.92	.88	.83
Small towns (< 100,000)	1.00	.95	.90	.85	.79
Rural towns and villages87	.79	.71	.63	.55
Education (complete secondary):					
Less than secondary:					
State sector men85	.82	.78	.74	.69
State sector women85	.87	.88	.90	.92
Private sector men59	.62	.65	.69	.74
Private sector women59	.65	.74	.84	.99
Vocational:					
Men	1.00	.97	.94	.91	.87
Women	1.00	1.01	1.03	1.05	1.07
VUZ degree:					
Men	1.20	1.20	1.20	1.20	1.20
Women	1.20	1.24	1.28	1.33	1.39
Private sector:					
Men	1.60	1.60	1.60	1.60	1.60
Women	1.00	1.08	1.18	1.30	1.46

* Baseline is the expected income in January 1991 rubles of Muscovite 18-year-old males in the state sector with complete secondary

† The age effect is evaluated at age = 40 and incorporates age and age².

residence type, education, and sector of employment; in January 1992 she was expected to earn 66%, and so on.²⁴ The education effects incorporate the main effect of the relevant dummy variable, its change over time, its interactions with gender and sector, and their change over time. Bear in mind that the education effects capture only the partial effect of being in a particular education category, though the magnitude of these effects can vary by the sector and gender of the respondent. The continuous age effect is evaluated at 40 years of age.

Table 7 offers scant support for hypothesis 1 (market transition increases returns to education). Among men, the increased returns to educa-

²⁴ The relevant entries were calculated according to the formula $\exp[\beta_{\text{woman}} + (\beta_{\text{woman} \times \text{age}} * 4)]$.

TABLE 7

EXPECTED EARNINGS MULTIPLIERS BY DEMOGRAPHIC VARIABLES, SECTOR, CLASS,
AND BRANCH, FULL MODEL

	January 1991	January 1992	February 1993	June 1994	December 1995
Baseline	222.13	143.94	132.80	138.32	126.52
Woman77	.74	.71	.67	.63
Age (40 vs. 18)	1.35	1.29	1.23	1.17	1.09
Locality (Moscow and St. Petersburg):					
Other large cities	1.00	.98	.96	.94	.91
Medium and small cities	1.00	.96	.92	.87	.82
Small towns	1.00	.96	.93	.88	.84
Rural towns and villages87	.82	.77	.72	.66
Education (complete secondary):					
Less than secondary:					
State sector men85	.85	.85	.85	.85
State sector women85	.88	.92	.97	1.02
Private sector men55	.60	.67	.75	.86
Private sector women55	.63	.72	.85	1.03
Vocational:					
State sector men	1.00	1.00	1.00	1.00	1.00
State sector women	1.00	1.03	1.06	1.09	1.14
Private sector men73	.78	.84	.91	1.01
Private sector women73	.80	.88	.99	1.14
VUZ degree:					
State sector men	1.23	1.18	1.13	1.07	1.00
State sector women	1.23	1.23	1.23	1.23	1.24
Private sector men	1.23	1.12	1.02	.91	.79
Private sector women	1.23	1.17	1.11	1.05	.98
Private sector:					
Men	1.63	1.63	1.63	1.63	1.63
Women	1.00	1.10	1.22	1.37	1.58
Branch (industry):					
Agriculture	1.00	.91	.83	.74	.64
Services90	.93	.96	1.00	1.06
Cognitive87	.83	.80	.75	.71
State/finance	1.00	1.00	1.00	1.00	1.00
Other	1.00	.93	.85	.77	.69
Class (VIIa. semi/unskilled in industry):					
Ia/IIa. managers	1.23	1.36	1.52	1.72	2.01
Ib/IIb. professionals:					
State sector	1.00	1.08	1.18	1.30	1.47
Private sector	1.00	1.03	1.06	1.10	1.15
IIIa. upper routine nonmanual	1.00	1.09	1.20	1.33	1.51
IIIb. lower routine nonmanual85	.85	.85	.85	.85
IV. proprietors	1.46	1.67	1.94	2.29	2.80
V. technicians/foremen88	.96	1.05	1.16	1.31
VI. skilled manual:					
State sector	1.14	1.14	1.14	1.14	1.14
Private sector	1.14	1.07	1.00	.92	.83
VIIb. semi- and unskilled, agriculture	1.00	1.00	1.00	1.00	1.00

tion materialized only for those with complete secondary schooling, and only in comparison to those with vocational training and those with less than secondary schooling employed in the state sector. In the private sector, men with secondary schooling lost ground relative to men with less than secondary schooling. Women did experience increased returns to higher education, consistent with hypothesis 1; however, women with complete secondary lost ground to women with vocational or less than secondary education. The only differences in returns to education within the state and private sectors pertain to the least educated group. The only sense in which hypothesis 1B is confirmed is that the negative return to having little or no education was greater in the private sector. But even this pattern evaporated over time, since the negative returns to "less than secondary" diminished for all but state sector men. Contrary to hypothesis 1C, men and women in the private sector with less than secondary education *gained* on those with complete secondary over time. With few exceptions—the most important being for women with higher education—market transition has not increased the returns to education in Russia.

Contrary to hypothesis 2A, the total gender gap has *increased* over time.²⁵ Contrary to hypothesis 2B, the gender gap has been *greater* in the private sector, since men's returns to private employment have exceeded women's. The gender gap has narrowed in the private sector over time: women's earnings rose from 43% of men's in January 1991 to 53% in December 1995. Yet it remains greater than in the state sector.²⁶ Finally, there is little evidence for hypothesis 2C that women's and men's returns to education are converging in the private sector—if anything, they diverge increasingly over time in both sectors. Contrary to the optimistic scenario of market transition theory, women fare no better in the Russian market than they did under the state socialist system, and in many cases they fare worse.

Proprietors gain relative to unskilled workers and other classes as market transition advances, confirming hypothesis 3. In fact, proprietors are

²⁵ The multipliers associated with "woman" pertain only to women with complete secondary employed in the state sector, due to the woman-by-sector and woman-by-education interactions. The "woman" multiplier derived from model 2, which omits the woman-by-sector interaction, is a somewhat better indicator of the overall trend; it declines from .65 in January 1991 to .54 in December 1995.

²⁶ These figures apply to men and women with complete secondary education. They are obtained by combining the appropriate main and interaction effect multipliers from table 7. When evaluating the gender-by-education interactions, the main effects must be taken into account. Thus, the gender earnings gap remained fairly stable among VUZ-educated Russians during the period under consideration: women's increased returns to higher education were almost wholly offset by the growing main effect of gender.

the only class whose expected earnings net of other variables grew during the reform period. Clearly, market institutions benefit entrepreneurs, as market transition theory predicts. The surprise is how few Russians have taken up self-employment.

The remaining hypotheses concerning patterns in class stratification are all contradicted by the data. Managers gain relative to workers, but their earnings in the private sector are no greater or less than in the state sector. Managers profit from their new-found freedom from ministerial control without having to pay consequences for poor enterprise performance: the breakdown of the administrative system permitted state managers to consolidate their positions within their organizations, while in the private sector ownership is too diffuse to serve as an effective check on managerial authority.

Professionals in the state sector have made some gains relative to unskilled workers—by December 1995 their expected earnings, net of other variables, are 47% higher. This finding supports hypothesis 5A. But the modest gain relative to unskilled workers is dwarfed by professionals' loss of ground to proprietors and managers. Moreover, within the private sector, where professionals should be best placed to accrue greater returns on their human capital, they are barely better paid than unskilled workers, even as market transition advances. Hypothesis 5B and hypothesis 5C are contradicted.

Skilled manual workers fare no better. The ones in the state sector do not gain on unskilled manual workers, and skilled manual workers in the private sector actually lose ground over time. Hypotheses 6A and 6B are disconfirmed.

In short, market transition in Russia has rewarded entrepreneurship and managerial authority, but it has not enhanced the returns to human capital. If anything, human capital has diminished in importance as a source of income advantages. Market transition theory predicts otherwise. The theory puts heavy emphasis on productivity gains driven by unleashed human capital potential.

The decrease in returns to human capital associated with market transition in Russia is even more evident from the partial returns to education when class and branch are controlled. Within each class category, the returns to higher education (relative to secondary) diminish over time for all but women employed in the state sector. The decline is most rapid in the private sector. Furthermore, in the private sector, the educational advantages of complete secondary education evaporate over time. Although initially the returns to education at these levels were indeed greater in the private sector, by December 1995 returns in the state and nonstate sectors are indistinguishable and modest, even negative. Market transition

theory hypothesizes greater and increasing income differentiation by educational criteria in the private sector. The opposite pattern holds in Russia.²⁷

Turning briefly to other effects in table 7, we see that the addition of class and branch variables lowers the gender earnings gap somewhat (compare the main effect of "woman" in tables 6 and 7), but not nearly enough to eliminate it. One source of the earnings disadvantage faced by Russian women is their concentration in the professional and lower routine nonmanual classes and underrepresentation among managers and proprietors (see table 2). However, a substantial residual gap remains once these compositional differences are controlled.²⁸

Younger Russians lucky enough to find employment—remember that between a quarter and a third of them are unemployed—have gained so much on their elders during the transition that the experience effect is overtaken by the cohort replacement effect. Age remained a highly significant net determinant of earnings, although not as a marker for experience, as it usually is in Western-style economies. The trend monotonically favored younger workers. Even in the late Soviet period, Russian workers experienced "age turnaround" while they were relatively young—43 years of age by our estimate (net of everything else in the model). The age turnaround has come ever earlier in recent years, falling to 33 years in 1996.²⁹ In most industrial societies, this turning point is reached around age 55 when the dominant experience effect is overtaken by reduced hours and the onset of retirement.

Income stratification based on economic branch and geography are often considered peculiar traits of the Soviet era (Zaslavsky 1982; Flakierski 1993). Market transition theory has not addressed them directly, so we do not list any predictions about branch and place. Nonetheless, as peculiarities of Soviet central planning, it seems reasonable to expect them to disappear after a period of reform. On the contrary, branch and geography

²⁷ One *AJS* reviewer suggested that opportunities to enhance income through second and third jobs may be greater for Russians with higher levels of education. If this were true, our results would understate the returns to education. We tested this possibility as best we could with the data at hand and found no evidence that taking secondary employment into account alters our findings (see appendix for details).

²⁸ This result would indicate that Russian employers discriminate against their female workers, if we knew that men and women worked equal numbers of hours. But we cannot be sure that the premise is true, so we cannot be sure that discrimination is the cause of this disparity.

²⁹ The peak is the age at which the partial effect of age becomes zero. That value is the age at which the partial derivative of log-earnings with respect to age (dy/da) becomes zero. Because $dy/da = \beta_a + 2\beta_{aa}a$, the value is equal to $1/2(\beta_a/\beta_{aa})$, where β_a is the regression coefficient for age and β_{aa} is the coefficient for age-squared.

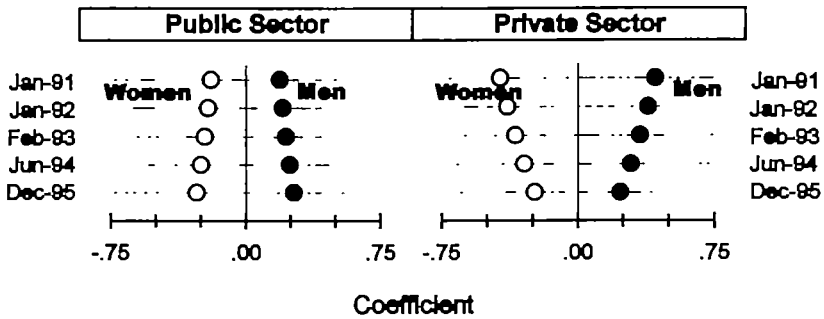


FIG. 2.—Effect of gender on monthly earnings by economic sector and time period, Russia 1991–95.

grew in importance during the market transition period in Russia. Services have gained on industry, the administrative arms of government kept pace, but the cognitive branches—education, health, science, and the arts—and agriculture lost considerable ground. Market forces are magnifying the peculiar geography of privilege in Russia (Zaslavsky 1982). Controlling for the other variables in model 6, only residents of rural areas earned less than Moscow and St. Petersburg residents in 1991. By the end of 1995, inhabitants of Russia's capital cities enjoyed a considerable advantage over all others, including 50% greater earnings than rural dwellers.

Figures 2, 3, and 4 summarize the differences in earnings by education, occupation, sector, gender, and time period. Men and women still earn sharply different wages even after many other factors are controlled. The returns to education barely rise. The gaps among occupational classes widen, but not in the way anticipated by market transition theory. The proprietors and managers reap big returns. But proprietors are a tiny sliver of the Russian labor force. Managers exploit their organizational positions to get rich at the expense of their workers. The other classes have lost in real terms.

DISCUSSION

Inflation has eaten away at Russians' earnings throughout the first half of the 1990s—beginning with the initial shock therapy in January 1992 and continuing through the first months of 1996. The real earnings of all but the top 10% of Russians have failed to keep up with hyperinflation. Only Russians in the upper decile of the income distribution have prospered (some greatly, some modestly). Measures of earnings inequality have risen as a result. By early 1995 the earnings of the lowest-paid Russian

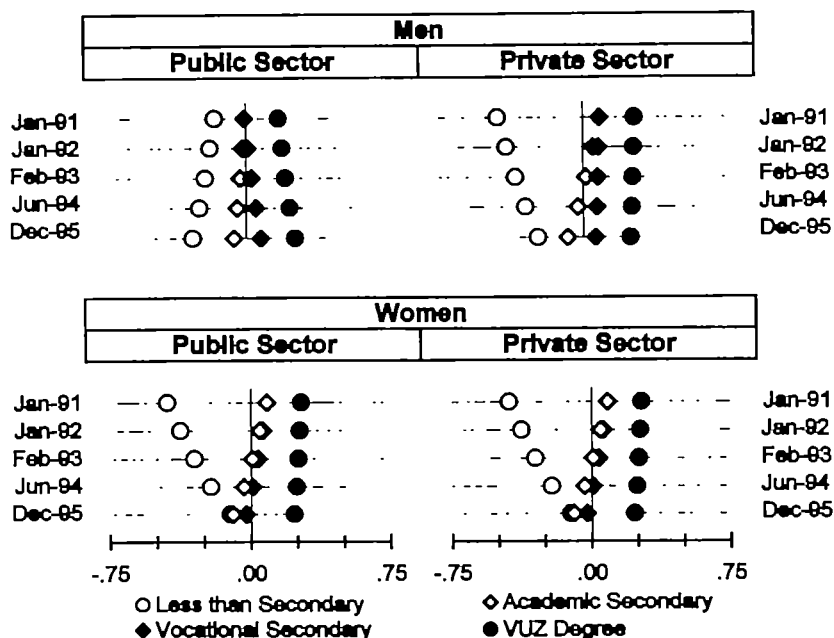


FIG. 3.—Effect of education on monthly earnings by sector and time period, Russia, 1991–95.

workers had hit bottom; their earnings were well below what prereform observers considered to be the subsistence level and could fall no farther. As the incomes of workers in the second earnings quintile fell, commonly used measures of inequality reversed and showed a slight egalitarian turn between 1995 and 1996. Unemployment and private sector employment have both grown sharply since the demise of the Soviet Union. To our surprise, the ranks of proprietors have not expanded.

Broadly speaking, only proprietors have been immune to falling real wages. Managers and upper nonmanual workers failed to keep pace with inflation but gained in relative terms as they came closer to matching the inflation rate than other classes did. The service branches and the youngest workers fared better than workers in manufacturing and those over 40 years old. Russians in other classes, particularly professionals in the cognitive and agricultural branches, lost relative earnings position. Radical market reforms have rewarded those classes already at the top of the earnings hierarchy and those in the new trade and service industries. Reforms have proven particularly devastating to the countryside.

Thus far, market transition in Russia rewards old-fashioned productive

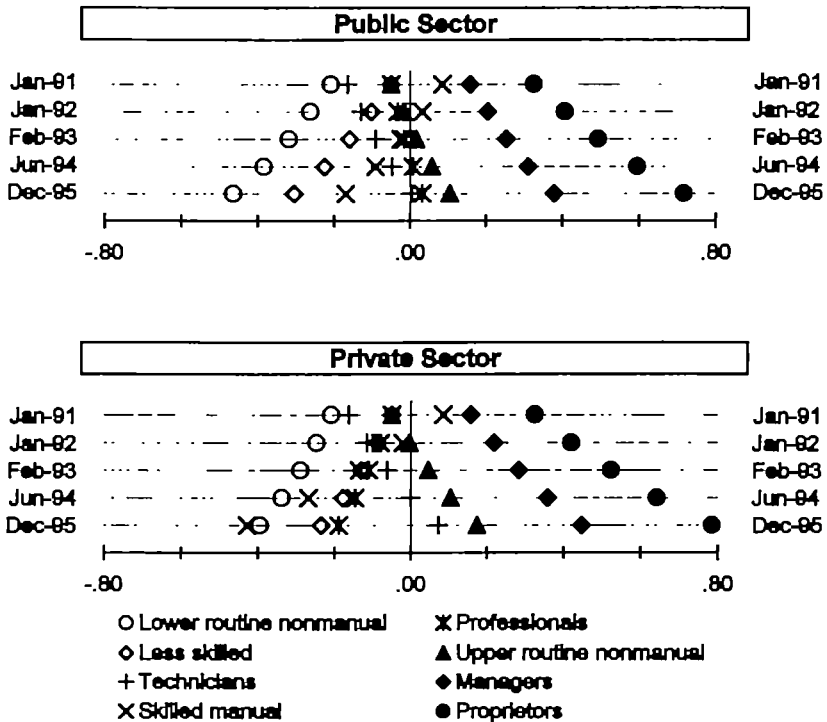


FIG. 4.—Effect of occupation on monthly earnings by sector and time period, Russia, 1991–95.

capital and managerial authority—not human capital. Contrary to expectation, the market reforms have not increased returns to human capital or “skill assets” (Wright 1985) relative to “organization assets” or productive capital. Professionals, along with skilled and unskilled manual workers, routine nonmanual, and technicians, must be counted among the losers rather than the winners of the market transition in Russia. Nor have women experienced any improvements relative to men.

Many formerly state socialist societies have gone through shock therapy. Could we have expected these disruptive outcomes? The leading theory of marketization—Nee’s market transition theory—boldly claims that the emergence and consolidation of market institutions in a state socialist system will reform the stratification system. Ownership, education, expertise, and manual skill will increase efficiency and productivity. People with those assets will reap rewards. Party membership, bureaucratic positioning, and gender will lose their influence. Our findings show that only one of these consequences—enhanced incomes for proprietors—has oc-

curred in the first five years of Russia's transition toward the market. Nee unequivocally predicts enhanced returns to education, but these have not materialized in Russia, where professionals must be counted among the losers rather than the winners of the market transition. Nor has marketization bridged the gender gap in earnings. Managers have reaped important rewards, which may be interpreted as supporting or countering the logic of market transition theory. In sum, we derived 16 predictions from market transition theory. The data contradict 14 of them. Russia's market reforms confound the expectations based on Nee's theory.

Why is the theory so wrong for Russia? Like any theory, market transition theory makes assumptions. Those assumptions may be wrong in the Russian case. A key assumption is that the lessons of China apply to Russia. But the Russian path to the onset of market transition differs from the Chinese path in many ways. China eased into a market transition. Each increment to GDP was rewarded with more economic liberalization. Most observers agree that each wave of liberalization further added to the already robust economic performance. Even in China, though, the precise mechanisms might be more "marketlike" than pure (Guthrie 1997). That is, the state's decisions to support some firms and "cut loose" others produce outcomes that are indistinguishable from what one would expect from market processes without any market actually being involved. All the while China enforced political stability by suppressing popular political movements, religion, and civil society.

Russia experienced less economic growth and more political change. The Russian economy lagged through the late 1970s and 1980s. This bred political instability and ultimately the collapse of the Soviet political system. Only then did market reform begin. What followed was economic shock therapy. The immediate outcome was steep contraction, rather than economic growth. While China is selling manufactured goods to the United States and Japan, Russian industry is producing less than it did 15 years ago. Russia's imports increased, and trade and consumer services have proffered opportunities to some and incentives to many more. But the centralized Soviet system left in its wake a tattered distribution system, endemic regional and local imbalances, a bloated and inefficient industrial sector, and a legal system most would-be investors find unreliable.

Compounding these problems, Russia has yet to achieve political stability. Far more democratic than China, Russia is nonetheless (or consequently) politically unpredictable.

These economic, legal, and political uncertainties discourage investment in productive enterprise and encourage the pursuit of short-term windfall profits in trade and consumer services. Manufacturers continue to bargain with suppliers, customers, and the state. In short, Russian capitalism is "merchant capitalism" (Burawoy and Krotov 1992; Burawoy

1997), characterized not by efficiency-oriented production that would reward human capital for its productivity but by a chaotic (if resourceful) scramble to capitalize immediately on what one has to sell. "The mass of the population falls back into a primitive, intensified domestic economy of dachas and subsistence plots and trading and buying their wares in the thriving bazaar" (Burawoy 1997, p. 1442). Supply and barter networks, access to locally scarce goods, connections with customs officials and local politicians, skill in the military and pleasure-providing arts—these, not education, are the most important forms of capital in the new Russian market. Numerous anecdotes regarding research scientists turning to prostitution or doctors engaging in smuggling (e.g., McDaniel 1996) provide graphic testimony that accords with our statistical findings. Market transition in Russia thus far has failed to reward education.

Our research shows that market transition is not a generic process. Predictions spawned in one case study do not travel well to another venue. Within China, "the conditions of marketization and the governance of firms in the new markets are . . . contingent on the state institutional structure and a firm's position in the hierarchy of the former command economy" (Guthrie 1997, p. 1298). So it is with entire societies too. The impact of market transition on stratification cannot be deduced from abstract, ideal-typical conceptions of the nature of market institutions. Rather, the precise institutional forms and opportunity structures that emerge with market transition matter—in a given firm or in a given country. The impact of market transition on stratification depends on sequence and timing of events—on the path taken (Walder 1996). In particular it depends on the preexisting class structure and how classes respond to reform. The persistence of managerial privilege despite reform in Russia stems from managers' strategic position in the transfer of ownership.

Russia differs from China in both the path into market transition and in the form marketization took once it was under way. We might speculate that Russia's experiences reflect a "shock therapy in response to decline" model. Stark's (1996) work on "recombinant property" and Szelenyi and Kostello's (1996) emphasis on how upper management "converts public property into its own private wealth" (p. 1094) show how patterns we observe in Russia occur in Europe, too. Comparative work focused there might yield new general theory. Or it might further reinforce the conclusion that history matters.

If the links between markets and stratification in formerly state socialist countries depend on the path taken out of state socialism and into capitalism, then taken-for-granted aspects of long-established market systems need a second look, too. Not only do functionalism and human capital theory have no answer for Russian professionals who wonder with dismay why the market has failed to provide them with the same rewards enjoyed

by their counterparts in advanced capitalist economies. They also have difficulty accounting for anomalies in Western Europe and North America. The rate of return on higher education fell in the United States in the early 1970s; its rebound contributed to the growth in income inequality over the past 20 years (Danziger and Gottschalk 1996). Fluctuations in fundamental parameters of the stratification process reflect the interaction of markets and politics that give societies their "design" (Fischer et al. 1996). To be predictive, theory must account for the political processes that ultimately determine the extent and character of socioeconomic stratification (Esping-Andersen 1994; Fligstein 1996).

APPENDIX

In constructing our variables, we aimed for consistency across instruments. Here we provide not an exhaustive account, but a brief outline, with some attention to ambiguities. Our *education* variable is a four-category recoding of responses corresponding to stages in the Russian educational system. Our "less than secondary" category includes respondents without complete secondary and no vocational education. We place respondents whose highest level of education was received in a vocational school (*professional'no-tekhnicheskoe uchilishche*, or PTU) into the "vocational" category whether or not they received a secondary diploma from their PTU. PTU graduates with secondary diplomas most likely have more in common with diplomaless PTU graduates than with students who have general secondary diplomas (see Connor 1991). General secondary, specialized secondary (*tekhnikum*), and some college are all recoded as secondary diploma. Our *sex* and *age* variables are straightforward recodes, in some cases requiring a transformation of year born. The aggregated cohorts in table 1 refer to age at the time of the survey.

Residence type necessitated combining variables on place of residence to identify Moscow and St. Petersburg residents and rural dwellers as well as population size of residence to distinguish among the three remaining categories based on city size. The CCSCP contained only the name of the current place of residence, coded in a scheme that distinguished rural settlements from urban. This allowed us to identify the first and fifth of our categories. The remaining cities were classified based on a listing of all Russian cities with populations over 100,000 and their population sizes (*Narodnoe Khoziaistvo Rossiiskoi Federatsii* 1992, pp. 87–91). The other data sets contained sufficient information to construct our categories by combination.

Workforce status in the CCSCP combined response categories to the question, "What were you doing last week?" We treated "at work" and "with a job but not at work" as employed; "looking for work" as unem-

ployed; "keeping house," "going to school," "unable to work" and "other" as not in labor force; and "on pension" as pensioner. The ISSP categories pertained to "current employment status" and distinguished among full-time work, part-time work, and "temporary, casual jobs"—all of which we treated as employed. Otherwise, the categories were the same. The omnibus categories contained five employed categories: "work at a regular job;" "work on contracts or work agreements;" "engaged in private business;" "individual work activity;" and "military service." They did not include an "unable to work" category.

In the CCSCP respondents were coded as employed in the state sector if they said their current place of employment was in the state sector of a collective farm (*kokhox*), private if they indicated the private sector. Russian ISSP respondents were asked, "Is your company, organization a public or a private one?" Their answers were coded as "state sector" or "public firm"—both recoded as "state"—or "private firm"—recoded accordingly. The omnibus data sets classified respondents' firms/organizations of employment into 10 categories, not all of which are unambiguously in the state or private sector. We opted to treat these categories as the state sector: state-owned, state joint-stock company, collective farm, consumer co-operative, public organization (foundation, political party, movement). These were coded as the private sector: individual labor activity, rented company, worker-owned company, private joint-stock company, private company, partly foreign-owned company. We coded the small number of "other" responses as missing.

Economic branch (also known as "industry") categories were imposed by the ISSP and omnibus data. Here are the original categories from the ISSP:

1. Industry, construction, transport, communication
2. Agriculture, forestry, hunting
3. Trade, catering, buildings maintenance, service sector
4. Kindergartens, school system, institutions of culture, arts, science, mass media, health care, system of higher education, sports
5. Voluntary nonprofit organizations
6. Public administration, financial, insurance bodies, public organizations
7. Other branch

The omnibus categories 1–4 and 7 were the same. However, Omnibus category 5 was "public administration, social organizations (parties, trade unions, funds, etc.), credit-finance sphere, insurance" and category 6 was "army, militia, Ministry of the Interior, Ministry of State Security." We were therefore compelled to combine categories 5 and 6, since they ap-

peared to be different splits of essentially the same categories. The CCSCP data included two-digit International Labor Organization (ILO) industry codes, which we recoded into the six-category classification based on the ISSP and Omnibus categories.

We operationalized our *class* schema based on respondent's occupation plus additional information to distinguish proprietors, who were not specifically identified in occupational classifications. The CCSCP employed three-digit ILO occupation codes, the remaining data sets used the four-digit 1988 International Standard Classification of Occupations (ISCO) classification. We mapped the respective classifications onto our class schema (precise mapping for class and branch codes is available from the authors upon request). We then identified proprietors. In the CCSCP we treat as proprietors only those who said their firm was in the private sector *and* said it was a cooperative (*kooperativ*). Proprietors in the ISSP are those who say they are self-employed *and* meet one of these additional criteria: 1988 ISCO code of main occupation is 1210 or response to a secondary question about main occupation (translated in the codebook as "what is your position?") of "businessman" or "senior manager." The Omnibus data sets allowed for what is probably our most reliable operationalization. Respondents were asked directly: "For your principle type of [work] activity, are you the owner or co-owner of a private business or enterprise—or do you work individually, by individual orders—or are you the hired employee of a state or private enterprise?" Respondents who indicated they were "owners or co-owners" *and* who described their firm of employment as a "privately owned firm" were coded as proprietors. We treat those who are not currently employed as "missing" on the class, sector, and branch variables.

Sampling weights.—It appears from discrepancies in bivariate distributions of education and gender across VTsIOM samples that VTsIOM did not fit the multivariate distributions of the four demographic variables in deriving the weights. Ideally, weights should insure that each cell of the four-way cross classification of demographic variables in each sample is proportionate to the corresponding cell for population. Instead, it appears that marginals were used to adjust the distribution of each variable separately, then multiplied. That the sample fluctuations in weighted marginals are nonetheless relatively minor suggests that the sampling variations in bivariate and higher-order interactions of the demographic variables are not severe. To derive weights for the CCSCP data, we fit the four-way education by gender by cohort by rural/urban proportions from the ISSP data set and adjusted the residence type distinctions not captured by the dummy variable by multiplying through the residence marginal weights.

Dependent Variable Specification and Model Estimation

We focus on earnings from primary occupation since we have data on these measures in each survey; therefore, we can test hypotheses about changes over time. However, we also have data on secondary employment and earnings from second jobs for the March 1993, July 1994, and January 1996 data sets. We looked at these variables to determine whether incorporating secondary incomes might affect our results—in particular, whether professionals have greater opportunities to enhance income through second jobs. We found that in each survey (i.e., at each point in time) professionals were more likely than average to have a second job. In January 1996, for example, 28% of professionals versus 21% of all employed persons reported a second job. Surprisingly, professionals earned less than average in their second jobs in March 1993 and July 1994, mitigating the effect of enhanced participation in secondary employment.

We also replicated the results we report in table A2 replacing primary-job earnings with total earnings (in log form) as the dependent variable (which is equivalent to primary-job earnings for all those who were interviewed in the first two surveys and all those who reported no second-job earnings in the third, fourth, and fifth surveys). Adding in these second-job earnings alters only one significant finding: the negative interaction effect of $VUZ \times Private \times Time$, which is significant when only primary-job earnings are the dependent variable, becomes insignificant in the analysis of total earnings. This revision does not affect our substantive conclusions since the two-way interaction of $VUZ \times Private$ remains significantly less than zero and market transition theory implies that both the two-way and three-way interaction effects ought to be positive. We also note that the coefficient for being a resident of Moscow is bigger in the analysis of total earnings; we read this outcome as an indication that Moscow offers more opportunities than other locales. We do not include these results in the article, but they are available from the first author upon request. We prefer to report the analysis of primary-job earnings because the definition of this variable does not change over time.

Our efforts to detect heteroscedasticity (prompted by comments from one of the *AJS* reviewers) revealed two sources of differential error variance. The standard deviation of the errors under model 6 varies significantly with time: .433, .577, .705, .680, and .615 for each survey (in chronological order). This pattern tracks the trend in the standard deviation of observed log-earnings, suggesting that unmeasured factors are key components of the trends in inequality. We also observe that the standard error of the residuals is greater in the private sector than in the state sector,

TABLE A1

INCOMES FROM PRIMARY OCCUPATIONS: RUSSIA

	January 1991	January 1992	February 1993	June 1994	December 1995
Classes:					
N	1,369	1,178	2,288	1,546	1,152
Ia/IIa. managers	351	233	435	333	214
Ib/IIb. professionals	273	159	181	154	123
IIIa. upper routine nonmanual	242	179	187	240	193
IIIb. lower routine nonmanual	166	116	152	119	94
IV. proprietors	555	718	713	514	590
V. technicians and foremen	223	161	138	125	99
VI. skilled manual	300	191	167	158	140
VIIa. semi- and unskilled, industry	273	185	155	142	118
VIIb. semi- and unskilled, agriculture	237	145	72	105	49
Highest/lowest	3.34	6.17	9.97	4.92	12.11
Professionals/managers78	.68	.42	.46	.57
Professionals/proprietors49	.22	.25	.30	.21
Professionals/semi- and unskilled, industry	1.00	.86	1.17	1.08	1.04
Managers/semi- and unskilled, industry	1.29	1.26	2.81	2.34	1.81
Proprietors/semi- and unskilled, industry	2.04	3.88	4.61	3.62	4.99
Sector:					
N	1,369	1,222	2,303	1,570	1,193
State	252	167	156	139	111
Private	512	311	358	334	222
Private/state	2.04	1.87	2.30	2.40	1.99
Economic branch:					
N	1,387	1,206	2,309	1,581	1,195
Industry ^a	284	210	188	203	143
Agriculture ^b	281	141	120	99	74
Services ^c	224	155	294	227	191
Cognitive ^d	223	117	151	117	91
State and financial ^e	237	173	221	166	165
Other	302	331	179	270	107
Agriculture/industry99	.67	.64	.49	.52
Services/industry79	.74	1.57	1.12	1.34
Cognitive/industry78	.56	.80	.58	.64
State and financial/services	1.06	1.11	.75	.73	.86
Education:					
N	1,387	1,289	2,309	1,581	1,198
Less than secondary	217	139	118	143	91
Vocational (PTU and FZO)	281	205	172	169	140
Complete secondary	257	164	204	162	130

TABLE A1 (Continued)

	January 1991	January 1992	February 1993	June 1994	December 1995
VUZ degree	341	230	261	192	173
VUZ degree/secondary	1.33	1.40	1.28	1.18	1.33
VUZ degree/vocational	1.21	1.12	1.51	1.13	1.23
VUZ degree/less than secondary	1.57	1.65	2.21	1.34	1.91
Sex:					
N	1,387	1,289	2,309	1,581	1,198
Men	323	212	242	228	160
Women	204	133	135	121	105
Women/men63	.63	.56	.53	.65
Age:					
N	1,387	1,289	2,309	1,581	1,198
18-25	219	135	181	271	141
26-40	284	190	217	170	138
41-55	259	178	168	158	131
Over 55	232	135	141	105	90
(18-25)/(26-40)77	.71	.83	1.60	1.02
(26-40)/(41-55)	1.09	1.07	1.30	1.07	1.05
(41-55)/Over 55	1.11	1.32	1.19	1.50	1.45
Residence type:					
N	1,387	1,289	2,308	1,581	1,198
Moscow and St. Petersburg	277	186	239	260	200
Other large cities (> 500,000)	266	197	206	166	155
Medium and small cities					
(100,000-500,000)	263	186	162	182	146
Small towns (< 100,000)	269	189	244	195	120
Rural towns and villages	247	126	135	116	86
Rural towns and villages/Moscow					
and St. Petersburg89	.68	.56	.45	.43
Other large cities/Moscow and					
St. Petersburg96	1.06	.86	.64	.78

* Includes transport, construction, communications.

* Includes other primary.

* Includes domestic, consumer, trade, insurance.

* Includes education, health, science, culture.

* Includes state organs, social organizations, finance.

.820 and .588, respectively. We replicated the results of models 2, 4, and 6 in table A2 using WLS instead of OLS (with weights varying by sector). The WLS results differ only slightly from the OLS results. A few *t*-ratios are smaller in the WLS results than in the OLS results, but WLS and OLS both agree on which variables are significant and which are not. We present the OLS results in this article because they are more familiar and not misleading in any way. The WLS results are available from the first author upon request.

TABLE A2

OLS MODELS OF RUSSIANS' MONTHLY EARNINGS FROM PRIMARY JOB, 1991-95

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Intercept	5.331*	5.428*	5.243*	5.331*	5.349*	5.403*
Date (January 1991)						
January 1992	-.311*	-.352*	-.336*	-.372*	-.415*	-.434*
February 1993	-.263*	-.342*	-.332*	-.396*	-.463*	-.514*
June 1994	-.076	-.201*	-.178*	-.280*	-.390*	-.474*
December 1995	-.025	-.206*	-.158	-.308*	-.442*	-.563*
Woman	-.413*	-.426*	-.340*	-.382*	-.215*	-.260*
Woman \times T	-.045*	-.039*	-.050*	-.033*	-.061*	-.040*
Age-18	.033*	.029*	.035*	.035*	.029*	.025*
Age-18 \times T	-.005*	-.003*	-.005*	-.005*	-.004*	-.002*
Age-18 ² ($\times 100$)	-.069*	-.059*	-.071*	-.072*	-.061*	-.050*
Age-18 ² \times T ($\times 100$)	.005		.006*	.006*	.005	
Locality (Moscow/St. Petersburg)						
Other large cities	.056		.060		.048	
Other large cities \times T	-.050*	-.034*	-.038*	-.019*	-.033*	-.020*
Medium and small cities	.047		.055		.044	
Medium and small cities \times T	-.067*	-.054*	-.055*	-.038*	-.053*	-.040*
Small towns	.087*		.094*		.049	
Small towns \times T	-.093*	-.068*	-.077*	-.048*	-.050*	-.037*
Rural towns and villages	-.083	-.141*	-.071	-.138*	-.124*	-.139*
Rural towns and villages \times T	-.136*	-.120*	-.118*	-.097*	-.059*	-.058*
Education (complete secondary)						
Less than secondary	-.196*	-.178*	-.143*	-.159*	-.157*	-.164*
Less than secondary \times T	-.042*	-.044*	-.050*	-.043*	-.014	
Vocational	-.003		.040		.007	
Vocational \times T	-.037*	-.036*	-.046*	-.030*	-.014	
VUZ degree	.256*	.190*	.265*	.183*	.294*	.204*
VUZ degree \times T	-.030*		-.031		-.073*	-.041*

Woman by education.				
Woman × less than secondary	.044	-.011	.021	
Woman × less than secondary × T	.052*	.066*	.039*	.039*
Woman × vocational	-.089	-.113	-.080	
Woman × vocational × T	.082*	.091*	.065*	.027*
Woman × VUZ degree	-.091	-.110	-.177*	
Woman × VUZ degree × T	.067*	.073*	.105*	.043*
Private sector		.672*	.437*	.491*
Private sector × T		-.064*	.010	
Private sector by education and woman:				
Two-way interactions:				
Private sector × less than secondary		-.547*	-.439*	-.431*
Private sector × less than secondary × T		.142*	.102*	.091*
Private sector × vocational		-.376*	-.390*	-.318*
Private sector × vocational × T		.116	.105*	.067*
Private sector × VUZ degree		-.238	-.217	
Private sector × VUZ degree × T		.061	.018	-.051*
Private sector × woman		-.664*	-.501*	-.490*
Private sector × woman × T		.152*	.104*	.095*
Three-way interactions:				
Private sector × less than secondary × woman		.094	-.091	
Private sector × less than secondary × woman × T		-.025	.021	
Private sector × vocational × woman		.189	.152	
Private sector × vocational × woman × T		-.112	-.074	
Private sector × VUZ degree × woman		.401	.334	
Private sector × VUZ degree × woman × T		-.126	-.114	

TABLE A2 (Continued)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Class (VIIa. semi- and unskilled, industry):						
Ia/IIa. managers					.181*	.206*
Ia/IIa × T					.104*	.101*
Ib/IIb. professionals					-.009	
Ib/IIb × T					.080*	.079*
IIIa. upper routine nonmanual					-.105	
IIIa × T					.124*	.086*
IIIb. lower routine nonmanual					-.216*	-.163*
IIIb × T					.026	
IV. proprietors					.476*	.381*
IV × T					.113*	.134*
V. technicians and foremen					-.155*	-.124*
V × T					.095*	.081*
VI. skilled manual					.087*	.133*
VI × T					.018	
VIIb. semi- and unskilled, agriculture					.019	
VIIb × T					-.018	
Branch (industry):						
Agriculture					.038	
Agriculture × T					-.101*	-.091*

REFERENCES

- Anurin, V. F. 1995. "Ekonomicheskaia stratifikatsiia: Attitudy i stereotipy soznaniia." *Sotsiologicheskie Issledovaniia* 1:104-15.
- Aslund, Anders. 1995. *How Russia Became a Market Economy*. Washington, D.C.: Brookings.
- Bian, Yanjie, and John R. Logan. 1996. "Market Transition and the Persistence of Power." *American Sociological Review* 61:739-57.
- Blasi, Joseph R., Maya Kroumova, and Douglas Kruse. 1997. *Kremlin Capitalism: Privatizing the Russian Economy*. Ithaca, N.Y.: Institute for Labor Relations Press.
- Boycko, Maxim, Andrei Shleifer, and Robert Vishny. 1995. *Privatizing Russia*. Cambridge, Mass.: MIT Press.
- Brada, Joseph C. 1996. "Privatization Is Transition—or Is It?" *Journal of Economic Perspectives* 10:67-86.
- Burawoy, Michael. 1997. "The Soviet Descent into Capitalism." *American Journal of Sociology* 102:1430-44.
- Burawoy, Michael, and Pavel Krotov. 1992. "The Soviet Transition from Socialism to Capitalism: Worker Control and Economic Bargaining in the Wood Industry." *American Sociological Review* 57:16-36.
- Burawoy, Michael, and Janos Lukacs. 1992. *The Radiant Past: Ideology and Reality in Hungary's Road to Capitalism*. Chicago: University of Chicago Press.
- Charles, Maria. 1992. "Gender Segregation of Occupations." *American Sociological Review* 57:483-502.
- Connor, Walter D. 1979. *Socialism, Politics, and Equality*. New York: Columbia University Press.
- . 1991. *The Accidental Proletariat: Workers, Politics, and Crisis in Gorbachev's Russia*. New York: Columbia University Press.
- Danziger, Sheldon, and Peter B. Gottschalk. 1996. *America Unequal*. Cambridge, Mass.: Harvard University Press.
- Erikson, Robert, and John H. Goldthorpe. 1992. *The Constant Flux*. Oxford: Clarendon Press.
- Esping-Andersen, Gösta. 1994. *Three Worlds of Welfare Capitalism*. Princeton, N.J.: Princeton University Press.
- Fischer, Claude S., Michael Hout, Martín Sánchez-Jankowski, Samuel R. Lucas, Ann Swidler, and Kim Voss. 1996. *Inequality by Design: Cracking the Bell Curve Myth*. Princeton, N.J.: Princeton University Press.
- Flakierski, Henryk. 1993. *Income Inequalities in the Former Soviet Union and Its Republics*. Armonk, N.Y.: M. E. Sharpe.
- Fligstein, Neil. 1996. "The Economic Sociology of the Transitions from Socialism." *American Journal of Sociology* 101:1074-81.
- Gerber, Theodore P. 1995. *In Search of the Soviet Middle Class: Scientists and Other Professionals in Post-Stalin Russia*. Ph.D. dissertation. University of California, Berkeley, Department of Sociology.
- Gerber, Theodore P., and Michael Hout. 1995. "Educational Stratification in Russia during the Soviet Period." *American Journal of Sociology* 101:611-60.
- Goldman, Marshall. 1994. *Lost Opportunity: Why Economic Reforms in Russia Have Not Worked*. New York: Norton.
- Gradolph, Rebecca. 1994. "Gender Stratification in Soviet Russia, 1991: Inequality of Opportunity and Outcome." Paper presented at the World Congress of Sociology, Bielefeld, Germany.
- Guthrie, Douglas. 1997. "Between Markets and Politics: Organizational Responses to Reform in China." *American Journal of Sociology* 102:1258-1304.

- Hout, Michael. 1996. "Inequality by Design: Myths, Data, and Politics." Working Paper no. 104. Russell Sage Foundation, New York.
- Hout, Michael, Erik Olin Wright, and Martín Sánchez-Jankowski. 1992. *Class Structure and Consciousness in Russia* (MRDF). Berkeley, Calif.: Survey Research Center.
- Intriligator, Michael D. 1994. "Privatisation in Russia Has Led to Criminalisation." *Australian Economic Review*, 2d quarter, pp. 4–14.
- Jones, Anthony, and William Moskoff. 1991. *Ko-ops: The Rebirth of Entrepreneurship in the Soviet Union*. Bloomington: Indiana University Press.
- Khakhulina, Ludmila A., and Milan Tucek. 1996. "Income Distribution: The Poor and the Rich in Postsocialist Societies (Some Results of a Comparative Analysis)." *Sociological Research* 35 (January–February): 20–32.
- Kharkhordin, Oleg, and Theodore P. Gerber. 1994. "Russian Directors' Business Ethic: A Study of Industrial Enterprises in St. Petersburg, 1993." *Europe-Asia Studies* 47:1075–1107.
- Kryzhanovskaya, Olga. 1995. "The New Russian Millionaires: Sources of Wealth, Style of Life, Political Views." *Sociological Research* 34 (May–June): 32–36.
- Mare, Robert D. 1995. "Changes in Educational Attainment and School Enrollment." Pp. 155–214 in *State of the Union*, vol. 1. Edited by Reynolds Farley. New York: Russell Sage Foundation.
- McDaniel, Tim. 1996. *The Agony of the Russian Idea*. Princeton, N.J.: Princeton University Press.
- Mikhalev, Vladimir. 1996. "Social Security in Russia under Economic Transformation." *Europe-Asia Studies* 48:5–25.
- Mroz, Thomas, and Barry M. Popkin. 1995. "Poverty and the Economic Transition in the Russian Federation." *Economic Development and Cultural Change* 22:2–31.
- Murrell, Peter. 1996. "How Far Has the Transition Progressed?" *Journal of Economic Perspectives* 10:25–44.
- Narodnoe Khoziaistvo Rossiiskoi Federatsii*. 1992. Moskva: Goskomstat Rossii.
- Nee, Victor. 1989. "A Theory of Market Transition: From Redistribution to Markets in State Socialism." *American Sociological Review* 54:663–81.
- . 1991. "Social Inequalities in Reforming State Socialism: Between Redistribution and Markets in China." *American Sociological Review* 55:267–82.
- . 1996. "The Emergence of a Market Society: Changing Mechanisms of Stratification in China." *American Journal of Sociology* 101:908–49.
- Nee, Victor, and Yang Cao. 1996. "Stratification in Path Dependent Societal Transformation." Paper presented at the 1996 summer meetings of the International Sociological Association Research Committee on Social Stratification, Ann Arbor, Mich., August.
- Nee, Victor, and Rebecca Matthews. 1996. "Market Transition and Societal Transformation in Reforming State Socialism." *Annual Review of Sociology* 22:401–35.
- OECD (Organization for Economic Cooperation and Development). 1995. *OECD Economic Surveys: The Russian Federation, 1995*. Paris: Centre for Cooperation with the Economies in Transition.
- Ospov, G. V., V. K. Levashov, and V. V. Lokosov. 1995. *Sotsial'naya i sotsial'no-politicheskaya situatsiia v rossii: Analiz i prognoz*. Moskva: Rossiiskaya Akademiia Nauk.
- Parish, William L., and Ethan Michelson. 1996. "Politics and Markets: Dual Transformations." *American Journal of Sociology* 101:1042–59.
- Parkin, Frank. 1972. "System Contradiction and Political Transformation." *European Journal of Sociology* 13:45–62.
- Perova, I. T. 1996. "Vnutrennii rynek truda: Zaniatost' v otsenkakh rabotnikov raz-

- nykh professional'no-statusnykh grupp." *Ekonomicheskie i Sotsial'nye Peremeny: Monitoring Obshchestvennogo Mneniia* 6:33-36.
- Róna-Tas, Ákos. 1994. "The First Shall Be Last? Entrepreneurship and Communist Cadres in the Transition from Socialism." *American Journal of Sociology* 100:40-69.
- Rutland, Peter. 1994. "Privatisation in Russia: One Step Forward, Two Steps Back?" *Europe-Asia Studies* 46:1109-32.
- Sachs, Jeffrey. 1992. "The Economic Transformation of Eastern Europe: The Case of Poland." *Economics of Planning* 25:5-19.
- Samodorov, Alexander T. 1992. "Transition, Poverty, and Inequality in Russia." *International Labour Review* 131:335-53.
- Smeeding, Timothy, and Peter B. Gottschalk. 1996. "Inequality in the Income Distributions of Modern Economies." Paper presented at the meetings of the Population Association of America, New Orleans, May.
- "Sotsial'no-ekonomicheskiĭ monitoring polozheniĭ rossiĭ I otdel'nykh stran mira." 1995. *Voprosy Statistiki*, no. 11, pp. 55-57.
- "Sotsial'no-ekonomicheskoe polozhenie rossiiskoi federatsii v 1992 godu." 1993. *Ekonomicheskii obzor*, no.1. Goskomstat Rossiĭ.
- "Sotsial'no-ekonomicheskoe polozhenie rossiiskoi federatsii v ianvare-tiale 1993 goda." 1993. *Ekonomicheskii obzor*, no.8. Goskomstat Rossiĭ Moskva: Respublikanskii Informatsionno-izdatel'skii tsentr.
- "Sotsial'no-ekonomicheskoe polozhenie rossiĭ v pervom polugodii 1994 goda." 1994. *Voprosy Statistiki*, no.1, pp. 53-68.
- "Sotsial'no-ekonomicheskoe polozhenie rossiĭ v pervom polugodii 1995 goda." 1996. *Voprosy Statistiki*, no.1, p. 44. Moskva: Respublikanskii Informatsionno-izdatel'skii tsentr.
- "Sotsial'no-ekonomicheskoe razvitiie Rossiĭ v 1995 godu." 1996. *Mir Rossiĭ* 5.3-93.
- Stark, David. 1996. "Recombinant Property in East European Capitalism." *American Journal of Sociology* 101:993-1027.
- Szelenyi, Ivan, and Eric Kostello. 1996. "The Market Transition Debate: Toward a Synthesis?" *American Journal of Sociology*, 101:1082-96.
- Walder, Andrew. 1995a. "Career Mobility and the Communist Political Order." *American Sociological Review* 60:309-28.
- . 1995b. "Local Governments as Industrial Firms: An Organizational Analysis of China's Transitional Economy." *American Journal of Sociology* 101:263-301.
- . 1996. "Markets and Inequality in Transitional Economies: Toward Testable Theories." *American Journal of Sociology* 101:1060-73.
- Webster, Lella M. 1994. "Newly Privatized Russian Enterprises." World Bank Technical Paper, no. 241. Washington, D.C.: World Bank.
- Wright, Erik Olin. 1985. *Classes*. London: Verso.
- Yanowitch, Murray. 1977. *Social and Economic Inequality in the Soviet Union: Six Studies*. White Plains, N.Y.: M. E. Sharpe.
- Xie, Yu, and Emily Hannum. 1996. "Regional Variation in Earnings Inequality in Reform-Era Urban China." *American Journal of Sociology* 101:950-92.
- Zaslavskaja, Tat'iana. 1995. "The Real Incomes of Russians Seen through the Prism of Social Assessments." *Sociological Research* 34 (July-August): 6-39.
- Zaslavsky, Viktor. 1982. *The Neo-Stalinist State*. Armonk, N.Y.: M. E. Sharpe.

A Frame Analysis of Favor Seeking in the Renaissance: Agency, Networks, and Political Culture¹

Paul D. McLean
University of Chicago

Actors invoke and manipulate diverse frames of meaning by assembling cues, taken from linguistic forms laid down in the cultural background, to build their relations with others. This article examines the use of these frames in Renaissance patronage-seeking letters, both quantitatively (through multidimensional scaling) and qualitatively (using discourse analytical concepts), to present an interactionist approach to the presentation of self and, in turn, to political culture. Writing strategies are only modestly predictable on the basis of actors' place in social structure, as actors write from achieved network positions and constantly aim to improve their position, maximize leverage, and build careers through letter writing.

I shall tell you, therefore, first, of what means I made use in order to become an intimate and follower of Gian Galeazzo, the duke of Milan; then I shall tell you how I went about winning the good will of Ladislas, king of Naples; finally I shall recount to you what sort of conduct enabled me to preserve the favor and good will of Pope Giovanni. I think, too, you will be pleased to learn of my various and different devices, my devious and seldom-used means, which have rarely been described. These are most useful ways to deal with men in civic life; therefore listen well to me. (Alberti 1969, p. 252)

I certainly agree, Lionardo, that the things Piero said all seemed to be wise and sound and full of prudence. His subtlety and considerable artistry were also clear to me. . . . But it seems to me that I want some other sort of thread and texture in talk on this subject. . . . You know the truth—how can anyone dream that mere simplicity and goodness will get him friends, or even acquaintances not actually harmful and annoying? . . . The world is amply supplied with fraudulent, false, perfidious, bold, audacious, and rapacious men. Everything in the world is profoundly unsure. . . . To deal with human wickedness in all its boldness, daring, and greed one must be

¹ I am grateful to Steve Ellingson, Alex Hrycak, Steve Laymon, Orville Lee, Susan Liebell, David Mandell, John Padgett, members of the Organizations and State-Building Workshop of the University of Chicago, and the *AJS* reviewers for their thoughtful and constructive criticism.

able to remain constant, temperate, and full of inner strength. These are the qualities I would like to see actually practiced by a man whose friendship I hoped to gain and enjoy. . . . So there it is; I want to be instructed in this sense concerning friendship, how to obtain it, increase it, diminish it, recover it, and make it permanent. (1969, pp. 264–67)

These passages from Leon Battista Alberti's treatise, *I libri della famiglia*, written in Florence in the 1430s, pose a perennial problem of social action: how to link strategic agency to cultural meaning or culturally valued ends. In the first passage, Piero Alberti outlines the appropriate forms of behavior and outward signs of honor by which to seduce men of various ranks in life into "friendship." The logic he expresses to construct personal networks and build a career is instrumentalist: choose the optimal interpersonal practices as means for achieving selfishly desirable ends. Friendship here is merely a means to self-advancement, and Piero provides a set of recipes for fabricating different friendships.

Yet as we see from the second passage, Adovardo Alberti clearly finds this set of recipes problematic, for two reasons. First, Adovardo wants a truer friendship, one based on a relationship with someone who really possesses virtuous qualities. Friendship—a normative commitment and a culturally shaped value—is an end in itself. Secondly, however, Adovardo is all too aware of human duplicity, and Piero's recipes seem insufficient, perhaps even for obtaining the mechanical forms of friendship he desires. In short, the dilemma is that one must use culturally sanctioned means of gaining advancement. But because these means often breed suspicion about one's own and others' sincerity and real interests, interpersonal solidarity begins to unravel. Somehow, practices subtler and less transparent than Piero's recipes are needed simultaneously to obtain an end we desire—friendship, group membership, social status—and to communicate to others that we "really" value what we say we do.² Adding to the dilemma is that achievement of a number of different relationships or network positions is not simply an additive process. We sometimes have to extricate ourselves from certain relationships we have achieved in order to gain leverage and to continue building a career in the manner Piero suggests (cf. Faulkner 1983), and very subtle tactics of decorum are necessary here, too.

Consequently, even as agents are undeniably "reflexive" implementers of culturally specific strategies (cf. Goffman 1974, p. 12; Clifford and Marcus 1986; Bourdieu 1990), culture must provide means for agents to get past this reflexivity to see each other's identity or deeper selfhood. How is the impression of a deeper self—the *public construction of a private*

² In fact, what we "really" value may be as much a product of what we achieve through interaction as something given prior to interaction. Thus, our interests become real for ourselves in a similar manner as they become real for others.

self—achieved? Utilizing Goffman's notion of frame analysis, I argue that this construction of self, which is so integral to achieving relationships, arises from the assembly of cultural cues and forms and from a publicly made ranking of relevant roles or motivations, provided in social interaction.¹ These constructions are also embedded in and refer to narratives of where these selves and their relationships have come from and where they are going (Linde 1987; Gergen and Gergen 1988; Steinmetz 1992).

In this article, I use both quantitative and qualitative techniques to analyze the presentation of self and the construction of careers in Renaissance Florence. A database of several hundred private letters through which Florentines sought favors from each other provides the raw material for this analysis. By demonstrating the relevance of framing and narrative to the embeddedness of actors in networks, I aim to articulate theoretically how the study of cultural practices must be integrated into a network-analytic approach to social reality (cf. Emirbayer and Goodwin 1994). Network ties, far from being autochthonous, are typically constructed through cultural agency: "stories are the essential vehicle for elaborating networks" (White 1992, p. 67). And among these forms of agency is the strategic, selective representation of the agent's network embeddedness in the course of interaction. This article aims to take us a step behind the macroanalysis of network structures (cf. Padgett and Ansell 1993) to examine the undertheorized efforts of individuals to occupy vacancies or capitalize on "structural hole"—like opportunities in the social structures to which they are connected (see Burt 1992). Examining this cultural work is necessary because it is the means by which social ties are created and maintained and individual mobility is enhanced. It is also the forum in which the nature of the obligations that these ties exert is determined.

More specifically, I offer a number of important cautions to practitioners of network analysis, deriving from both macro and micro concerns. First, different sorts of relationships will be shown to be constructed, at an aggregate level, employing different constructions of motive. This places in some doubt the simple stackability of different types of ties, for different ties may involve rather inconsistent presentations of self, potentially to the same alter. Saying this, however, it is also

¹ It can also come through violating rules of decorum. Thereby, the recipients of cultural signals locate identity precisely where the veneer of etiquette is punctured and infer the nature of the author behind that construction (Goffman 1974, chap. 8; Foucault 1975). This is as true for the Piero-like strategist who baldly states his crass, self-interested motives as for the Adovardo-like seeker of spiritual relationships. Rule-breaking behavior helps define the boundary between acceptable behavior and violations and thus sustains the boundary as a common point of contention (cf. Laitin and Wildavsky 1988).

clear that actors frequently try to bootstrap new relationships to already achieved ones whenever possible. Positions (whether achieved or still sought after) in different networks will have determinate impact on each other, mediated to a considerable extent by the narration of their connection (or disconnectedness) on the part of actors. Second, it will be shown that specific types of ties to particular alters—say a relationship of gaining an office through the graces of a particular patron—have very different meanings to different actors and are constructed through rather different cultural frames and symbols. Such a fact means we lump all such office-holding relationships into a single social network matrix at some peril to our understanding of reality. Third, it will be documented how even within a particular relationship, the precise substantive significance of the tie and its strength over time is subject to considerable variability. Just where actors are in their careers, for example, needs to be addressed to better estimate the meaning of a particular tie and its consequences. Further, in the course of constructing ties, actors implicitly negotiate the strength and exclusivity of that tie. Simply reading interpersonal loyalties off of network patterns ignores this negotiation to the detriment of understanding network dynamics. In short, network analysis must attend much more to how actors get into (and out of) relationships and to how those relationships are managed by cultural means. Failing this, the intensity, durability, flexibility, heterogeneity, and additivity of different social networks cannot be well understood.

FRAME ANALYSIS AND THE STUDY OF CULTURE

Although much attention has been devoted recently to matters of culture in sociology, we still need a conceptual framework and a set of methods for empirical research that show how culture “works” (Schudson 1989) by productively integrating the understanding of culture as a set of flexible rule-based practices (Cicourel 1973; Swidler 1986) and as a web of meanings in which agents obtain socially recognized identities for themselves (Geertz 1973). Strategic action and the construction of interests (the establishment of identities) *both* arise on the basis of the practical cultural tools or building blocks out of which social interaction is constructed. Several desiderata ideally should be satisfied by a good technique for studying culture: identifying common practices and their canonization in the form of etiquette, rules of decorum, and so on; tracing out empirically and theoretically why certain cultural practices are more robust than others; identifying departures from empirically common patterns and their implications; identifying what kinds of combinations of cultural practices seem most common or most persuasive; and locating linkages, interpenetra-

tions, and correspondences between the complex patterns of cultural structure and social structure.⁴

I use the notion of *frame* and associated analytical concepts (Goffman 1974; Snow et al. 1986; Sarat and Felstiner 1988; Tannen 1993; Gumperz 1992) to tackle some of these requirements. The notion of framing is useful because it spans instrumentalist and interpretivist connotations of culture, implying both negotiated meaning and the strategic erection of resemblances and exemplars relevant to constructing a particular context for action. Frames organize experience, but not in a simple unidirectional path between stipulated actions, nor in terms of simple binary oppositions (cf. Alexander and Smith 1993). Rather, culture is better understood as a multidimensional cognitive space in which a variety of frames are deployable. Then we can analyze "strips of activity" (in this case, Renaissance patronage letters) for keywords, images, and phrases that concretely construct, suggest, or evoke which frame of meaning is in effect at a given time. Of course, one needs to spell out which frames and cues are available to understand the cultural terrain upon which given instances of agency take place. Further, we will want to see how different frames can be juxtaposed since the strategic lamination of vocabularies of motive on top of each other can create an impression that certain motives are purer, deeper, more "real" (Goffman 1974, p. 2), or more representative of private interests and goals than are others.

This cultural agency in terms of framing typically involves particular techniques that are amenable to analysis. For example, we may look for how writers focus reader attention on certain events, characters, or perspectives. We can look for amplification of themes or frames through repeated keyword use (Snow et al. 1986), for strategies of politeness that help maintain face (Brown and Levinson 1987), and for ways of communicating an expectation of assistance. We can look for the inclusive and exclusive, boundary- or frame-creating use of pronouns as deictic expressions (Muhlhausler and Harré 1990) and for contextualization cues that communicate the proper context for understanding a message through the text of the message itself (Duranti and Goodwin 1992, p. 18; Gumperz 1992; Hanks 1992). In short, we should analyze strips of activity for the way they conjure background meanings and play with their location

⁴ Naturally, culture and structure, however analytically separable they are, will be interpenetrating, because individuals and institutions are positioned in both, cultural classifications are typically integral in defining spaces in social structure, and placement in different social spaces brings one into contact with different streams of agents and cultural practices. This viewpoint has been succinctly stated by Sewell (1992, p. 27; cf. Archer 1988).

vis-à-vis those background meanings as a means of situating the self of the agent in relation to the self of the alter and to the social identities of competitors for scarce goods.

THE QUANTITATIVE ANALYSIS OF POLITICAL CULTURE:
FLORENTINE PATRONAGE LETTERS AS A CORPUS OF
PATTERNED DISCOURSE

Renaissance Florentines sought each other out constantly, through face-to-face interaction or through letters, to obtain favors such as favorable tax treatment; offices for themselves, friends, relatives, or clients; writs of safe-conduct; and securing support in various disputes. As the volume of offices available increased with the conquest of the *contado* over the course of the later 14th century (Becker 1968), the volume of correspondence grew; and as the burden of taxation grew with the prosecution of wars in the early 15th century (Molho 1971), and as a smaller and smaller subset of the elite came to control appointments through patronage (Najemy 1982; Brucker 1977), seeking favor—or recognition of one's worthiness (Griffiths et al. 1987)—became a very urgent matter in a highly competitive environment. Letters, both private and official (and sometimes both simultaneously), were the lifeblood of the commune (Kent 1991; Najemy 1993). Examined in the aggregate, we find in them many regularities of expression, practical cultural building blocks canonized into a set of rhetorical tropes and cue words that signal frames of meaning by which individuals constructed and prosecuted their preferences in imitation of established practices (White 1978; Holland and Skinner 1987, p. 87). Thus, for simple expository purposes, we can treat the frequency and ubiquity of particular words and turns of phrase as measurable "cultural objects" (Griswold 1987).

Measuring the use of these objects across cases allows us to trace which cues and framing devices are most used, the purposes to which they are put, and whether they are used in conjunction with or in contrast to other ones. Thus, the ensemble or structure of frames that comprise quattrocento Florentine practical culture is empirically determined; conversely, assessing individual letters by means of which elements of the cultural "toolkit" they utilize places writers of letters in particular parts of the cultural space. Some of these building blocks were discussed in letter-writing manuals, and writing letters was considered a core element of the average Florentine merchant's son's curriculum (Grendler 1989; Weissman 1989; Morelli in Branca 1986). Once this terrain is mapped using frequency calculations and multidimensional scaling, a subsequent section will use some of the discourse analytic tools developed in the tradition of Goffman's work on frames to examine how particular agents assemble, key,

break, and laminate different frames of meaning, thus differentially appropriating these cultural objects as they erect public constructions of their private selves. The quantitative and qualitative elements of the analysis are essential for each other: we cannot know what instances of cultural practice are exceptional without a general overview of those practices, nor can we understand enough about agency and the practice of framing without examining particular letters for techniques of construction and self-presentation.

The database for this analysis comprises 869 letters, composed between roughly 1380 and 1460, by means of which Renaissance Florentines sought favor and friendship from influential patrons. The data is used in conjunction with a wealth of information about individual writers and receivers of patronage letters. Several thousand letters were read and considered for coding. The final sample is not fully representative, nor could it be, given that extant letters are preponderantly addressed to members of a few powerful families: the Medici, the Strozzi, the Sacchetti. Letters between family members are underrepresented since career-building dynamics tend to be much less observable there. Letters had to be legible—not a negligible consideration with this data, as Florentine scholars will attest. Finally, letters of modest length (typically no more than one full page) were targeted, with an eye toward selecting a subset of letters that would include (1) letters between both political allies and political enemies, (2) letters from social elites as well as from those of modest background, and (3) both isolated letters (the most common case) and, when possible, series of letters from the same writer. Consideration of the letters written between roughly 1421 and 1436 was quite comprehensive, and representation of these letters is quite high. This period was focal because information about the wealth, residence, and network embeddedness of individuals was richest for this period, in the aftermath of the 1427 *catasto* (Herlihy and Klapisch-Zuber 1985). The sampling is more precarious for later years, and further archival research is planned to redress this problem.

Some of these letters have been used by other researchers, most notably in Dale Kent's (1978) impressive work on the rise of the Medici in the 1420s, but they have not been subjected to the kind of quantitative analysis performed here, nor have they been discussed so extensively in terms of the discourse analytic tools used here. My particular focus is on how people do politics through writing. Where I summon the facts, I employ them instrumentally to understand the way in which discourse is embedded in, modified by, and constitutive of the structure of social relations and the stream of events.

Letters were assigned a primary content and were coded for the occurrence of keywords and phrases, approximately 50 in number, chosen for

how important they appeared to be in the background cognitive map of Florentine culture.⁵ The average occurrence of various keywords in different types of letters is reported in table 1.

These keywords are of two sorts. (1) Some act chiefly as substantive cues that suggest a framework of meaning in which the letter-as-encounter is taking place—Am I framing my competence or my relationship to you as based on my sense of honor (*onore*), on personal friendship (*amicizia*) or loyalty, on filial affection, on justifiable need (*bisogno*), on your generosity (*magnificenza*), on virtue or objective worth, or on the dictates of communal service? Which of these framings gets used subtly changes the cast of the relationship being constructed. (2) Some provide skeletal semantic forms upon which and through which a writer's personal case is presented in a recognizable and sanctioned format. These are turns of phrase that customarily accompany and rhetorically signal the introduction or reinforcement of a request: "In service to them, *work things out* as seems best to you with the Signoria that they obtain their wish, which is quite justified" (in loro servizio *adoperare* quello vi pare colla S. che obtengano la domanda che è assai giusta; MAP XII, 161); "Tell me what you wish and I will do it *willingly*" (se io posso fare qualcosa per voi m'avisate e faròllo *volentieri*; MAP XI, 456); "I am *certain* you cannot fail me" (*certo* Io ho fermissima speranza che non mi possa mancare; MAP XI, 432); and the ubiquitous "*I beg of you* as much as *I know how and am able to*" (*priegovi* quanto *so e posso*; C.S. III, 130:192). Words and phrases of this sort anticipate and animate the making of a request. What we might consider rhetorical inflation is in fact commonplace here, a fact that cannot be documented well without an extensive quantitative review of keywords.⁶

In addition, each letter was coded where possible for the age, income, and neighborhood of sender and recipient, plus the date of the first priorate (a measure of political prestige) and embeddedness in economic and marital networks of the senders' and recipients' families. For each case, a set of variables was calculated measuring the differences between sender and recipient in age, income, and political prestige, and the proximity of

⁵ Choice of keywords was guided specifically by the author's reading of numerous Renaissance treatises, *ricordanze*, public documents, and the like, and based on a sense, in the course of reading thousands of letters, of which sorts of syntactical devices and substantive framings recurred in a salient fashion.

⁶ Some of these turns of phrase have clear resonance with formulas used in the dictaminal tradition of letter writing in the late Middle Ages and early Renaissance. A whole literature exists on this subject, but for an introduction to it, see Murphy (1974), Kristeller (1983), Witt (1988), and McLean (1996, chap. 2). The primary sources (letters) used here are from the Archivio di Stato, Firenze, and include pieces from the Carte del Bene, Carte Stroziane, III^e serie (hereafter C.S. III), Conventi Soppressi, 78 (Conv. Soppr.), and Mediceo Avanti il Principato (MAP) collections.

TABLE 1
AVERAGE FREQUENCY OF SELECTED KEYWORDS BY CONTENT OF LETTER

Type	N	SUBSTANTIVE KEYWORDS						SEMANTIC KEYWORDS			
		Honor	Diligence	Virtue	Magnificence	Amicitia	Kindship	Adopters	Certain	Progress	Quanto to e posio
Assurance	43	.51	.23	.05	.16	.56	.56	.19	.40	.81	.16
Difference	17	.29	.24	.06	.18	.41	.59	.06	.24	1.47	.18
Flattery	12	1.17	.25	.00	.83	.17	.50	.17	.67	.75	.00
Information	70	.39	.17	.03	.19	.30	.31	.16	.14	.21	.04
Military news	96	.33	.16	.02	.11	.24	.09	.17	.16	.43	.02
Office	78	.83	.21	.15	.32	.41	.33	.68	.35	1.12	.13
Recommendation	107	.33	.20	.04	.25	.53	.27	.28	.14	.95	.18
Unspecified favor	46	.30	.26	.09	.07	.33	.63	.37	.22	1.20	.15
Property recovery	68	.19	.34	.01	.19	.22	.24	.22	.32	1.47	.18
Prison release	39	.26	.21	.10	.46	.51	.31	.44	.31	1.23	.10
Tax	35	.17	.63	.00	.09	.40	.23	.29	.20	1.03	.06
Thanks	32	.16	.31	.13	.06	.31	.34	.22	.28	.69	.25
Total	869	.35	.24	.05	.19	.38	.32	.29	.23	.89	.13

NOTE.—Each cell is calculated by the total number of incidents of that keyword divided by total number of letters of that type. Only data on the 10 most frequently found types of letters plus "flattery" and "difference" letters are reported here. Letters from these 12 categories contain 643 of the total number of letters collected (869).

their residence, since the text and framing of letters has more to do with the relative positioning of sender and recipient than with their absolute traits. The network variables are inherently relational. I use these variables later to assess the extent to which cultural practice is shaped by social structural positioning of different sorts.

The "assurance" letters in table 1 are those in which the writer primarily assures the recipient of his loyalty. Letters about *differenze* pertain to matters of dispute between the writer and someone else and often request some adjudication or mediation from the recipient. Much of the extant correspondence consists of simple informational letters and letters in which military engagements or military supply problems are discussed. "Office" letters are those in which the writer seeks office, usually but not always for himself, while recommendations are those in which the author typically recommends a friend, relative, or acquaintance for a specific favor or general goodwill. Requests for unspecified favors, requests for assistance in recovering property, and requests for assistance in gaining release from prison, whatever the offense (felonies, bankruptcy, hostage taking, and military capture all figure here), are also common subjects. "Tax" letters are those in which cash-strapped individuals sought loans or influence from rich or powerful others to meet their fiscal obligations. "Thanks" letters are those in which gratitude is expressed for past favors done. For purposes of simplification, the balance of this article will focus most on the features of office-seeking letters.

Each cell in table 1 reports the average occurrence of a particular keyword in a particular type of letter and can be compared to the average incidence of the given keyword across all letters printed on the "totals" line.⁷ By extension, reading across the rows gives a sense of the available keywords in the Florentine toolkit and their relative weighting for framing the average office letter or the average request for money letter, and so on. Few keywords appear once or close to once per letter, even when the data is segmented by type of request. No hardwired template existed that absolutely required certain kinds of requests to be framed in determinate ways.

Specifically, table 1 tells us that the notion of honor appears approximately two-and-a-half times as often in office letters as the average across all types and cases and is also used with frequency in letters flattering to the recipient and those expressing personal loyalty or attachment (assur-

⁷ Separate chi-square statistics were calculated for the presence/absence of each of the variables reported in this table by letter type. All were found to be significant at a $P < .05$ level except *virtù* and "certainty." Letter salutations invoking honor ("Honorable sir") or brotherhood ("Carissimo fratello"), the most formulaic of openings, are not significantly linked to particular types of letter.

ance). Building relationships with an eye to these ends strongly leads toward the framing of oneself or one's alter as being concerned with honor. The concept of honor was, and is, a fluid concept whose meaning changes situationally and over time (Pitt-Rivers 1977; Bourdieu 1977). In 13th-century Florence, honor was most attached perhaps to the notions of family solidarity and vendetta (Lansing 1991), a sense it sometimes retained into the 1400s. For various late 14th-century Florentine authors such as Lapo da Castiglionchio and Coluccio Salutati, honor meant typically a certain self-confident disposition and a particular lifestyle involving landed wealth and the pursuit of the highest, most adventurous mercantile vocations—in other words, the lifestyle of the post-Ciampi ruling elite. In the memoirs of ordinary merchants such as Paolo da Certaldo (Branca 1986), honor meant primarily an unbesmirched reputation in business dealings. But with Leonardo Bruni, chancellor of Florence in the early 1400s and a key figure in the development of civic humanism (Baron 1966), honor became firmly wedded to taking a role in civic administration. The regularity with which office was cast in terms of bestowing honor upon its occupant fits well within the Brunian conceptual landscape.⁸ Beyond these different substantive implications of being honorable however, honor seems to involve such notions as being one's own man, not running away from danger, and having a public reputation for integrity and for savvy, drawing upon each as the situation requires. In short, even as political matters emerged at the surface of how honor was pursued in the early 15th century, a variety of significances remained in the background. The resultant multivalency of the notion of honor had an important effect in making switching between frames of meaning more possible.

Table 1 also shows that petitions for office highlight the magnificence of the potential patron more than the average—the display of his power and generosity in bestowing a position, a code of action rather distinct from the more bourgeois and circumspect attitudes of Florence's merchants and ruling elite in the early quattrocento. Where flattery is the main aim of the letter, honor and magnificence may both be legitimately emphasized. By contrast, when the supplicant finds himself in undeniably dire straits, such as in prison, the unconditional generosity and potency of the patron implied by the magnificence framing appear appropriate and efficacious, whereas discussions of honor are typically considered out of place.

Amicizia—a form of friendship in Renaissance Florence in which senti-

⁸ There is even a semantic connection between *onore* (honor) and *onori* (offices). Singular and plural versions of the term, however, were carefully coded separately, and most often the context dictates that the singular form of the word connotes honor, integrity, or public respect, rather than "office."

ment and instrumentality were typically conflated (Trexler 1980; Silver 1989) and which often signified the existence of favoritism, political connections, or partisanship—is another important frame for pursuing the achievement of office. Nepotism in the distribution of offices was widespread, and expectations of action based on nepotism are surprisingly baldly expressed in some letters and implicitly invoked in others where the recipient was requested to “work things out” (*adoperare*) on behalf of the supplicant. *Amicizia* is commonly used in letters of “assurance” and as the rhetorical grease lubricating letters of recommendation, the main venue for the activation of weak tie networks (Granovetter 1973).

Neither honor, nor magnificence, nor *amicizia* figures very strongly in letters where some kind of economic relief was sought, whether tax relief or assistance in the recovery of property. Instead, individuals making these requests stress qualities like diligence (*diligenza*), effort (*fatica*), and ingenuity (*ingegno*) in the patron. Patrons are framed here by favor seekers as being adept at cutting through red tape, but there is less signaling in these letters of interpersonal commitment, reciprocity, or relationship building—an intriguing difference between politically and economically oriented interactions not commonly attended to in the clientelism literature. Finally, note how infrequently reference is made to *virtù*. Though both Dante and Machiavelli discuss this notion extensively (albeit in diametrically different senses), it is largely avoided in the everyday discourse studied here, which should make us wary of jumping to conclusions about which elite concepts and values frame everyday interaction. Where *virtù* does figure in the discussion of office holding, one senses it plays a subordinate role to other considerations.⁹

Table 1 thus provides a sense of how different sorts of request on average generate different sorts of framings of self and other. When we turn to individual letters, we may then come equipped with better grounded expectations concerning the kinds of terms the Florentines typically employed for seeking advantage and constructing their relationships.

What table 1 cannot do, however, is tell us if there is an “average” practice of actually framing a request that employs a composite of keywords that refer to and laminate together a variety of analytically separable frames of meaning, or if the “averages” are misleading aggregations across subsets of letters that actually employ mutually exclusive frames.

⁹ See, e.g., Cresci di Lorenzo de' Cresci's letter to Cosimo de' Medici, December 17, 1439 (MAP XI, 336), in which he wryly remarks concerning the decision of a certain group of officials to appoint Medici *amico* Antonio Martelli to a position in lieu of a different candidate, that “they came to this [choice] willingly on account of his *virtù*, and the contemplation of you” (a questo sono venuti volentieri per la virtù sua e per la contemplazione di te).

A better technique for getting at this question and for representing the rhetorical space in which letters get constructed is multidimensional scaling (MDS). MDS generates a spatial image of the distance keywords (and therefore frames) are from each other, distance being a function of how the keywords are distributed across the cases of letters. Cooccurring keywords are placed close together; keywords used in largely different subsets of letters are placed far apart.

Figure 1 plots selected keywords used specifically in office-seeking letters and highlights how honor, *amicisia*, expressions of objective need (*bisogno*), patronal magnificence, and clientelistic service constitute differentially employed frames for construing the significance of an office-based relationship between sender and recipient. Among the less frequently used keywords in the center of the plot, expressions of obligation (*obrigato*) and faith or trust (*fede*) are placed closest to expressions of service, while considerations of advantage (*utilità*) tend to appear more with requests set up in terms of the instrumentally and transactionally inclined framework of *amicisia*. Florentines saw different framings of relationships with others as conceptually distinct and as having distinct implications.

Framing oneself in terms of the pursuit of honor was to highlight one's autonomy, concern for values, and equal membership in an elite that shared a sense of the importance of office holding and obligation to the city. "Honor" meant one was willing and worthy to hold office on the basis of one's own character and abilities, not on the basis of one's loyalties. By contrast, stressing the magnificence of one's alter effectively reduces one's own claim to honor. The language used in this framing of self is highly deferential; the worthiness of the self disappears in the foregrounding of the power and glory of the alter. Letters adopting this framework make considerable use of *pregare* (to beg), the primary deferential verb of petitioning. Here too, we tend to find greater use of the notion of need (*bisogno*) as a justification for assistance: the man of honor craves a position; he does not, strictly speaking, "need" it. Further, need may be a justification for assistance independent of the interpersonal responsibilities of clients and patrons or fathers and sons and thus stands in contrast to the framing of oneself as a "servant."

The *amicisia* frame implicitly treats self and other as more or less equal—at least equally committed to a partisan project—and often engaged in an exchange of favors. In this aspect, it resembles the (contentiously) egalitarian quality of the honor framework. But it is distinguished from writing in terms of honor precisely in replacing objective considerations of merit with subjective considerations of loyalty. The tension involved in identifying oneself and alter as members of the same class or ruling elite through the use of the frame of honor, in order to obtain a desired end, despite the fact that self and alter are allied with different

parties or groups of *amici*, is precisely what produces the complexity of the correspondence between Ormanno degli Albizzi and Averardo de' Medici discussed at some length below.

Finally, the service or clientelistic framing tends to be deferential, like the framing of alter in terms of magnificence. But like the *amicisia* framework, it tries to draw a boundary of partisanship inclusively around self and other. In highlighting favors owed to one's followers, the "service" framing aims to constitute the relationship between self and alter as a hierarchical but affective relationship. Overall, then, these different frames calibrate the social distance between sender and receiver in different ways and ground the terms of their relationship and their joint treatment of others in substantially different ways. Hence, in figure 1, the horizontal axis measures differences in the representation of social status, while the vertical axis measures differences in the representation of degree of relationship. These culturally informed representations of social proximity and distance are integral to the construction of relationships.

Figure 2 maps these competing framings onto the sample of office-seeking letters to guide us toward understanding how actual letters work within this cultural terrain. Honor is the predominant (though not exclusive) frame for letters on the left side of the figure, while patronal magnificence figures strongly in letters on the right. Letters toward the top of figure 2 try to reinforce the relationship between writer and recipient in affective hierarchical terms (as, e.g., in case 65 [MAP III, 13], where Carlo Biliotti concludes by referring to himself as "vostro servidore fidellissimo"). Those toward the bottom build the request for office chiefly on the basis of *amicisia*, or favors owed between friends. This mapping is crude, but overall, the MDS plot of keyword use offers useful suggestions about different constructions and tones of self-presentation.

The question obviously arises about why each of these different available framings would get chosen by particular writers. The answer is, to a noteworthy extent, rooted in writers' placement in social structure—both status hierarchies and relational networks, as demonstrated in table 2. This table reports on the linkage between variation in language usage across cases and background social structural factors such as the relative age, wealth, and political prestige of letter recipients and senders, as well as the degree of their connectedness through marital and economic networks and the degree of neighborhood separation between them. The upper half of the table indicates that the difference between recipients and senders of letters found in the core of the MDS plot in figure 2, in terms of wealth and political capital and distance between their neighborhoods of residence, is smaller on average than it is for cases in the periphery. Senders here are also more likely to be from the same family as recipients, more likely to be linked to the family of letter recipients through direct

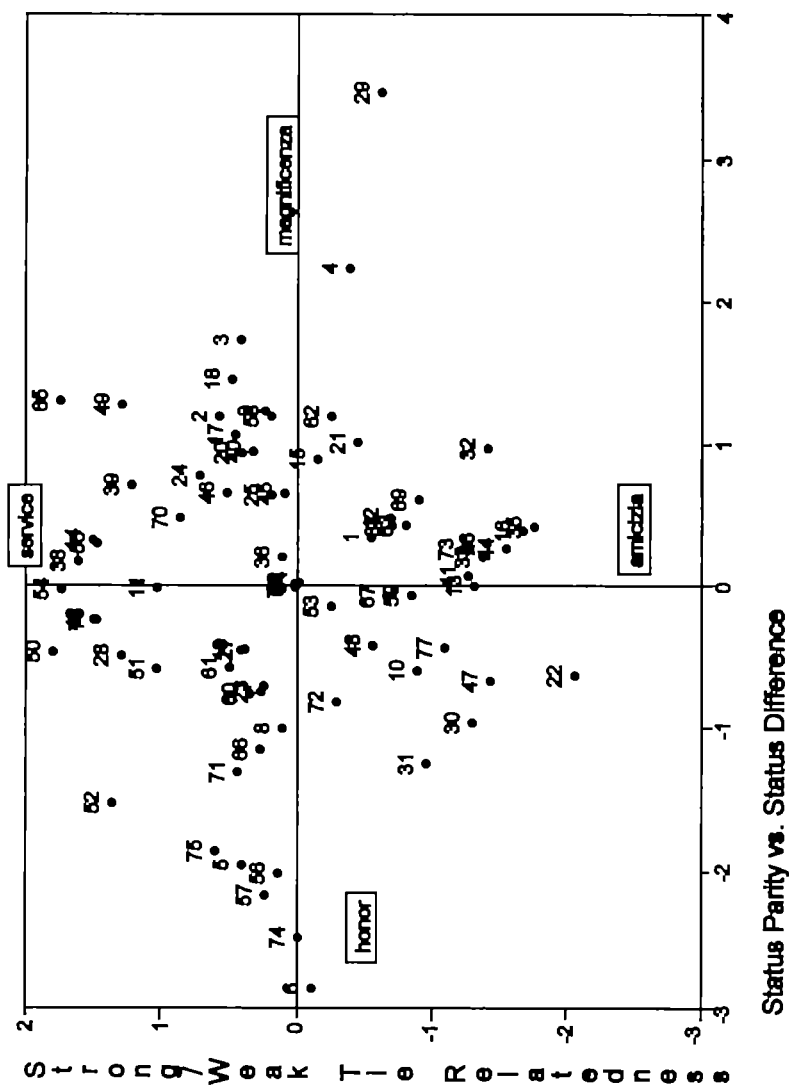


FIG. 2.—MDS plot of letters seeking office

TABLE 2
RELATIVE STATUS OF RECIPIENTS AND SENDERS OF OFFICE LETTERS AS PLOTTED IN FIGURE 1

CENTRAL CLUSTER LETTERS VERSUS ALL OUTLIERS AND OUTLIER SECTORS						
RECIPIENT/SENDER STATUS	Central (N = 31)	Outliers (N = 47)	Magnificence Sector (N = 10)	Service Sector (N = 13)	Honor Sector (N = 11)	Assault Sector (N = 13)
Age difference (in years)	4.71 (17)	4.19 (26)	-3.50 (2)	5.20 (10)	14.29 (7)	-5.14 (7)
Wealth difference*	1.00 (21)	1.04 (28)	1.67 (3)	1.11 (9)	-14 (7)	1.67 (9)
Political prestige difference†	.67 (27)	1.00 (36)	1.50 (6)	1.36 (11)	.38 (8)	.82 (11)
Proximity of residence‡	2.50 (30)	2.61 (44)	2.89 (9)	2.62 (13)	2.30 (10)	2.67 (12)
Path distance in marriage network§	2.52 (31)	3.32 (47)	5.20 (10)	2.77 (13)	3.09 (11)	2.62 (13)
No. of direct marriages¶	1.54 (28)	.71 (38)	.50 (6)	.73 (11)	.67 (9)	.83 (12)
% of letters within family	.26 (31)	.09 (47)	.00 (10)	.08 (13)	.09 (11)	.15 (13)
Gross economic relatedness¹	1.47 (19)	1.90 (31)	1.60 (5)	1.90 (10)	2.43 (7)	1.67 (9)

NOTE.—Nos. in parentheses are valid *n*s for the relevant variable.

Wealth difference score represents approximately an exponent of 10, 10⁶ being the factor by which recipient's income in 1427 outstrips that of sender.

†Political prestige in Florence was customarily measured in terms of the year one's patrilineage first placed a member in the city's priors. Larger numbers indicate a greater degree to which any member of recipient's family preceded any member of sender's family into the ranks of the city's most esteemed administrative body.

‡The residential relationship between recipient and sender was coded *1* if they lived in the same *parishes* of the city, *2* if in different *parishes* within the same quarter, and *3* if they lived in different quarters. Thus smaller numbers indicate greater proximity.

§Marriage network relatedness indicates average path distance between sender and recipient families within sectors using Pedgnet's (1994) marriage network data. Thus smaller numbers signify greater proximity in the Florentine marriage market.

¶Direct marriage ties equals number of times recipient's and sender's families intermarried in the 30 years prior to and one year after the date of the letter. Thus larger numbers betoken more underlying relationship between sender and recipient.

¹Gross economic relatedness is a recalibrated sum of all economic ties between sender's and recipient's families as reported in the 1427 *cedola*.

marriage ties, and to be more proximate in the Florentine marriage market. Thus, cultural innovation and emphasis on meaningful tropes appear to be generated more by social outsiders than by those at the core (cf. Hunt 1984), precisely because they need their letters to communicate what their social position in itself cannot.

The second half of the table divides the periphery into sectors. Given the mapping presented in figure 1, we should expect the letters in the "honor" sector to be written by supplicants of an age, income, and political prestige comparable to their alters, while the numbers for the *magnificenza* sector should be higher. In fact they are, except for age.¹⁰ Those using honor should be closer in the marriage market to their alters than those using the *magnificenza* framing, and this too appears, on average, to be the case. Comparing the service-oriented and *amicizia*-oriented sectors, we predictably find age differences and political prestige differences between recipient and sender to be greater in the former. Since *amicizia* in one of its primary senses is based on weak ties (Granovetter 1973), sender's and recipient's neighborhoods might well be farther apart for cases here than in other sectors; so, on average, they are. Marriage connections between the sender's and the recipient's families are clearly least likely for those who resort to deferential framing in terms of *magnificenza*. The greatest amount of economic connectedness between families holds for the sector in which status parity seems strongest, and is next strongest in the service sector, the strong tie/status difference sector, again as we might expect.

In the final analysis, considering the small sample size for the various sectors, these trends should be considered suggestive rather than conclusive. A more important caveat is in order, however, for predicting how letters are framed on the basis of the placement of actors in underlying social structures remains a risky business. There are a number of cases in which fairly low-status writers write in intimate terms to high-status recipients.¹¹ This intimacy across social distance is endemic to the nature

¹⁰ Age is an odd case here. Many of the letters in the honor sector were written by Ormanno degli Albizzi to Averardo de' Medici. The two came from families of comparable prestige, but Averardo was old enough to be Ormanno's father. I discuss the details of this correspondence extensively below.

¹¹ Cases 15, 37, and 59 in the core sector of fig. 2 and cases 13 and 57 in the *amicizia* sector are cases of this sort. In case 13 (MAP II, 201), the artisan Michelozzo di Bartolomeo Michelozzi writes confidently to Averardo de' Medici: "Many times I discussed with you whether it might happen that you could give some position to Giovanni my brother, were the occasion to arise, and now I believe you will have had one." A certain Matteo appointed as *sottoscrivano* on a galley the commune was equipping would be unable to take up his post, as he was in the field, and this would afford Averardo perfect opportunity to provide Giovanni with a post. Michelozzo continues:

of clientelism, as discussed by Eisenstadt and Roniger (1980, 1984) and others. The difficulty always lies in assessing where individuals lie at the present moment on their path toward achieving intimacy *despite* their apparent structural position based on static indicators, for the very point of the form of communication studied here is the dynamic construction and improvement of relationships.¹²

THE QUALITATIVE ANALYSIS OF POLITICAL CULTURE: PATRONAGE LETTERS AS EXERCISES IN FRAMING

In what follows, a variety of examples of different sorts of framing of motive and castings of relationships are presented, proceeding toward more complicated forms of laminating together different frames of meaning to produce textured public presentations of self. Throughout this section, I want to draw attention to the way individual writers use the substantive frames I have described so far, but also to how they construct their letters to achieve a portrait of themselves and their motivations by amplifying certain themes through repeated keyword use, communicating their confidence in being well served, gauging the limits of their politeness, linking their interests rhetorically to those of their would-be patrons, ranking their preferences, and providing the context for their requests in the course of their writing. These techniques for self-presentation and "networking" appear modern, even if they are harnessed to building relationships in substantive terms rather different from those we might seek.

Arguably, the simplest way to request assistance in obtaining office is to ask for it bluntly, as in a letter of Francesco di Niccolo Baldovinetti to Forese Sacchetti written October 12, 1409 (Conv. Soppr. 78, 324:172; case 61 in the core of fig. 2):

In this regard [Francesco had just indicated when his present occupation would conclude], I pray of you as much as I can, that if you were to see yourself able to provide me with some office or other commission, either keeping track of your soldiers or inspection or some similar thing, I pray you advise me of it, for it would be singularly pleasing to me. . . . And I pray

"I have it on authority from Ser Martino [another Medici partisan] that these matters rest in your hands, and thus I recommend this to you."

¹² In fact, a logistic regression procedure that aimed to predict whether or not letters would use formal language based on status and relational attributes misclassified a number of precisely those cases where low-status Medici benchmen write in intimate terms to their patrons. This difficulty precisely demarcates the treacherous waters where a Bourdieuan approach that treats social structure and cultural practice as homologous founders (cf. Bourdieu 1984). Low-status yet intimate correspondents could well display both valencies of their identity in the text and tone of their letters, something to look for in the qualitative analysis of interaction.

that you keep me in mind in whatever you might see that would increase my honor.¹³

This letter, written using the informal "tu," gives one the sense that the two men were already well acquainted with each other, so that hardly any framing of Baldovinetti's character or situation needed to be provided in the letter, nor any rationale for Sacchetti to take action on Baldovinetti's behalf. Of course, this in itself is a construction of the self-in-relation.

Sometimes this informality, however, takes the form of positive politeness (cf. Brown and Levinson 1987) and becomes an explicit part of the framing, as in another letter from the core of figure 2 (case 53), a letter from Marcello Strozzi to his kinsman Simone Strozzi (C.S. III, 131:12):

Speaking with such familiarity (*dimestichezza*) as we may together, I pray of you that you be the agent (*operatore*) whereby Giannozzo Cavalcanti, my brother-in-law, who has need of earning a little money, and also of having some office (*honor*), receive from you the office of inspector (*rassegnante*), so that this May he is sent out. I pray this of you considering it to be a singular pleasure of mine.

The opening clause here represents an effort by Marcello at reminding Simone of the intimacy they share and thus sets up and justifies the candor of the request that follows, as if to say: "I could not speak so frankly to someone whom I did not already consider my confidante." Further, note how different rationales for granting this favor—*parentado*, need, honor—are rhetorically linked by Marcello as all conducing to Simone taking action on Cavalcanti's behalf. The rhetorical representation of the multiplexity of common ties and motivations linking Marcello and Simone is intended by Marcello to deepen Simone's sense of commitment to him.

But there also exist slightly more complicated versions of framing requests for office in terms of *amicizia*. First, we have a letter of May 14, 1431 from Bartolo Ridolfi to Averardo de' Medici (MAP III, 126; case 22 in the *amicizia* sector of fig. 2) in which Bartolo writes on behalf of a certain Benedetto da San Miniato, "who is a person for whom I would always exert myself to do good, both because he merits it and I am obliged to him, and because he is a friend and relative of your, and my, dear friends."

Because Benedetto had decided not to return home, he was now in

¹³ "In questo mezzo ti priego quanto posso se vedessi potermi fare alcuna onore o alcuna comessione o descrivere vostri soldati o rassegnare o simile cosa ti priego ne sia avisato riputandomelo in singulare piacere . . . E priego mi abbia a mente in quello vedessi fusse di mio onore." Note the two distinct senses of the word "onore" in this letter (cf. n. 8 above).

search of a new position, and Bartolo was urging Averardo to help get him a position with a different *signore*. He went on to write:

For this reason if there is any favor or pleasure whatsoever you can do for me, as much as I am able I pray of you that it please you to fix him up with Signore Michele as his chancellor. . . . He is an able person in this and everything, and if you will do it, as I know you can, you will be many, many times thanked by Signore Michele; and to me, especially, among whatever pleasures you may afford me, this is one of the greatest ones. I will not express myself further on this, because I know it is not necessary, and because things will be arranged in the way you want, and your will and mine are the same thing.

The boundaries of friendship and the confluence of interests are actively constructed in this letter: yes, Benedetto is an able person, but more important, he is a friend of the friends of both of us, and so a circle of *amicizia* is drawn around him and Averardo. Bartolo urges Averardo to consider the consequences of providing this assistance, "speaking for another" (Schiffrin 1993) by anticipating for Averardo the kind of gratitude he could expect to receive from Signore Michele, and from Bartolo himself. Bartolo quite explicitly ranks the awarding of this favor against others that he might have in the past requested, that he currently was soliciting, or that he hypothetically may request in the future. Consequently, this request is contextualized and prioritized in light of other potential requests, intimating where Bartolo's deepest preferences lie. Further, he gives an account of how Benedetto has chosen not to go home as it would be a waste of time, "and in truth he speaks the truth, for you know all the offices there are rotten and the country is in darkness." This short narration of the reality of the situation gives greater justification to Benedetto's plan and Bartolo's portrayal of his need. Finally, Bartolo actively communicates an expectation of being assisted by claiming an identity between his and Averardo's will and a confidence that Averardo can bring to pass whatever he wishes to have done.

Under certain circumstances, the request for favor on behalf of friends had to be hedged with rhetorical concern for honor. A letter of Palla di Nofri Strozzi to his kinsman Simone di Filippo Strozzi of August 11, 1421 (C.S. III, 132:50; case 56 in the honor sector of fig. 2) demonstrates such a texturing of motivations:

I wrote to you praying that you contrive, within the bounds that honor requires, that Galeotto di ——— be elected to the constabulary of this place, telling you of his *virtù* and gentility, and also that he would be cognizant of his place as I understand it. And although I hear at the constabulary there is no place because the electors have deliberated, still the captain has not been provided. It would be dear to me if concerning this you might operate for my friend. May it not be bothersome to you to arrange the consent of Raffacane, who is one of the electors. And recommend him to him,

and contrive some action in this matter as much as you see possible, while retaining honor. He is a man of such *virtù* that I do not doubt that he would honor those who were to elect him, and who were to perform any favor for him.

Palla, scion of the largest and one of the most prestigious families in the city, offers this recommendation for his friend but does so in the context of honor: twice he emphasizes that pulling strings should be done within the demands of honor, and although Galeotto can be expected to be loyal to those who would help him, Palla tries to suggest this loyalty is a product of his sense of honor, as though returning favors were a matter of maintaining one's integrity and impeccable self-image rather than a matter of instrumental attachment to a group for the sake of individual advancement. This letter highlights the potential tension between acting with an eye to honor and acting with an eye to *amicisia*, and the expression *salvando l'onore*, which afforded the recipient the entitlement of saving face if he failed to satisfy a request, became a stock element in the way Florentines went about making requests and recommendations. Here the expression is particularly apt, for Palla is inviting Simone to intervene illegitimately in a political process that has already been decided. This letter therefore suggests that, even as kinship networks were an important instrument in the erection of political networks, the obligations that the latter networks entailed had considerable potential for straining or even rupturing kin relations.

The foregoing letters exemplify simple, emphatic, and then complex uses of the *amicisia* frame, the last being a case in which the friendship framing is complemented by or contrasted with other sorts of ties and motivations linking actors. These different manners of invoking *amicisia*, however, can be contrasted as a whole with examples in which supplicants portray themselves as servants of their patron. For instance, we have the letter of Francesco di Luigi Benintendi to Cosimo de' Medici of September 1, 1435 (MAP XII, 25; case 39 in the service sector of fig. 2):

My honorable elder, This letter only for the reason that I find myself in Pisa as you know, and I am searching for some occupation with which to sustain myself until the galleys leave, but still I have been able to find nothing—nor will I obtain anything without your help, such that I pray of you, as your faithful follower and servant, that in service to me you write a short note to the Marchese Malaspini, who is here as you know. . . . And further if you were to see another place in which to situate me with your letter, I pray of you, that wherever you put me, I will do you honor. Perhaps I use too many words, but the great faith I have in you moves me, such that I pray of you that I be recommended to you, as I am in a bad sort of state at present.

This letter, signed "your least servant," goes to considerable length to stress the dependence of the sender on Cosimo for help, as he claims im-

plicitly to have no other source of assistance. But it also implicitly suggests that this assistance is warranted on the basis of a standing relationship: I have been "your faithful follower and servant." Moreover, various cues are used to communicate simultaneously Cosimo's power and Benintendi's confidence in being helped. Rather than the letter communicating much in the way of news to Cosimo, Benintendi frames it as though Cosimo no doubt were already aware of his circumstances: as you know, I am in Pisa; as you know, the Marchese is here. These cues contextualize his request in terms of shared understandings with Cosimo. Rather than the strong framing betokening his desperation, Benintendi tries to suggest it is a function of his enthusiasm for Cosimo. As a result, Benintendi sets up his relationship to Cosimo as though assistance should flow automatically from the penning of the letter.

A letter of Francesco Nardi to "my" Averardo de' Medici similarly builds on the assumption that patron and supplicant have an ongoing relationship, and perhaps more explicitly than Benintendi's asserts the formative influence of the patron on the supplicant's capacities and character. Nardi sought a renewal of his term of office as *podestà* of Città di Castello (MAP II, 292; case 15 in the core of fig. 2):

Honored singularly like a father, In the past few days, having confidence in your paternity as I know I can, I wrote to you, that it please you to be with Cosimo, and that he deign to write to Rome that I might have a reappointment to this office, and from you I have had no response, so that I think there will have been some bad delivery with the mail, such that with this letter again I pray of you, and beg, my Averardo, that you might wish to be the *operatore* by means of whom I get this reappointment, or some other office, so that I might return there. It would give me the heart to be a man like the others as you intend me to be.

Failure of patrons to respond was often attributed to problems with the mail, whereas Averardo was daily inundated with requests, sometimes for identical positions, and no doubt had to determine which requests would get priority. Lack of confidence in obtaining satisfaction, however, was rarely voiced by supplicants. The last sentence here is particularly striking, for it offers an implicit, almost contractual commitment of loyalty in return for a favor, and moreover, explicitly links the development of Nardi's manhood—even his own self—to Averardo's actions.¹⁴

¹⁴ A letter from Andrea di Pagolo della Scarperia to Averardo on April 5, 1430 (MAP II, 328) is very similar. Andrea writes: "Please contrive (*adoperate*) that I may stay here, and let me thus be able to say that you were the one who made me the man I would want to be. . . . As you are all my hope, I recommend myself to you; contrive what is needful that I be reappointed." A diametrically different incentive for being granted a position is that provided by a certain Antonio on July 12, 1435 (C.S. III, 112:189) while looking for a position as a knight with Pazzino di Palla Strozzi, the new *podestà* of Rimini: "When I return to my office in Montemerlo, with your, and

Looking next to the *magnificenza* sector of Figure 2, we find a rather complicated letter from Luca Carducci to Giovanni di Cosimo de' Medici of February 3, 1454 (MAP I, 252; case 4):

Thinking of how many times we have bothered you, requesting of you favors (*servigi*) for us and for our friends, through the mediation of Antonio degli Strozzi your brother-in-law and our honorable brother in love, it is not without trepidation I thought to write to your Magnificent Lordship to ask a new favor (*gratia*) and service, but provided it is an appropriate thing, noble and generous souls are in their magnanimity no less ready to grant what anyone were to demand. Considering the *virtù* and benevolence of your Lordship, I presume with great confidence to ask of your Lordship a favor, which is that Matteo Carducci my brother, at present *podestà* di Piccioli, by your *virtù* and efforts (*operatione*) be reappointed for one year or for six months, for the election has been postponed for one year. . . . Such that to your Lordship in this instant I recommend this matter, hoping that thus Matteo will carry himself to better things, and always we will remain obligated to your magnificent Lordship, whom may God in His grace always maintain and prosper.

Note here the repeated addressing of the recipient as "Lordship," even though this is a letter between fellow Florentines. Thus it stresses, even more than the actual status difference between families would, the dependence of Luca on Giovanni for assistance.¹⁵ The heavy framing of Giovanni in terms of his nobility and magnificence accomplishes the goal of appearing to be highly deferential—this in spite of the intimate connection the recipient and sender both have with the same individual, a contextual fact that Luca himself provides for us in the body of the letter. The aim here seems to be to remind Giovanni of a personalistic claim to assistance, without explicitly relying on this criterion. In fact, the personalism inherent in this request is actively portrayed as somewhat distasteful to Carducci, so that the reasonableness of the request, or the reading of it as an opportunity for the exercise of magnanimous generosity, takes center stage.

Finally, a more explicit distinction between motivations occurs in a letter of May 22, 1434, from the notary, Iacopo di Antonio di Salvestro, to Matteo di Simone Strozzi (C.S. III, 112:175; case 48 in the core of fig. 2):

Honored like a father, I heard that Messer P[iero] your relative was elected captain of Cortona, such that I beg of you that, if he is not so furnished,

my, Smeraldo, we will have such beautiful parties at your house as we used to in the past." Antonio looks backward to a shared set of experiences, rather than forward to a promised new level of loyalty.

¹⁵ Thus, one may observe here, in the middle of the fifteenth century, a move toward more flowery and deferential forms of written interaction in imitation of princely rhetoric and with more self-consciousness about the *gentilezza* associated with verbal dexterity and rhetorical sophistication (cf. Trexler 1980; Kristeller 1983).

you speak to him about me being his notary. And if he wants information about my affairs, I went with Antonio di Tedice degli Albizzi to Pisa when he was *podestà*, and with Buonaccorso Pitti to Prato. I am not concerned with the salary so much as with acquiring friends. I say no more for the moment. May Christ always watch over you. And if I can do anything in Pistoia, I am ready to do whatever be pleasing to you.

In some respects, the writer here simply reproduces customary practices to present his request. For example, the letter uses a customary “*vi priego*” formula; the salutation partakes of both honor and fatherhood frames common in office letters; and it ends with a formulaic expression of generalized reciprocity that is typical in favor-seeking letters. The mention of friends suggests in a fairly standardized way Iacopo’s desire for reliable connections with an eye toward advancement and, ultimately, his desire for relationships.

However, we have moved beyond the Middle Ages world where standardized gestures or other objective correlates are adequately communicative of internal dispositions (cf. Hood 1990). How do we read through this letter to the person? How is the person’s motivation constructed in the letter? The various writing strategies we have observed so far—contextualization cues, speaking for others, emphasis, cajolery and assurances, circumlocutions, ranking of preferences, and so on—are all designed for constructing a believable representation of the person.

Two main techniques can be observed in this letter. First, note how this claim for friends is framed: “I am not concerned with the salary so much as with acquiring friends.” This is an answer to an unposed question: Which is more relevant to understanding Iacopo’s motives in obtaining this appointment, friendship or money? Iacopo’s writing contextualizes his request for favor ‘reflexively’ through the very text by means of which that request is made, assertively establishing the relevant frame for understanding his request. In this publicly constructed image of his private motivation, Iacopo claims his role in *amicisia* corresponds to his true self, his ‘real’ motives: “I am willing to trade off salary for friends.” And in Florence’s politically charged climate, to invoke *amicisia*—especially in regard to administrative office—was to evince a willingness to enter a world of personal loyalty and political dedication.

The second device pertains to the representation of social networks discursively. Actors are positioned in networks, in status groups, and in relationships. Yet in discourse, relevant past relationships can be highlighted, and problematic ones can be cast in shadow. As a notary, Iacopo was probably not wealthy, although he might have had opportunity to fraternize with powerful men, as notaries were a vital source of labor and knowledge for the commune. Lacking a surname, he lacked a significant

marker of social prestige. Hence, he invokes his connection to two prominent Florentine families. Like the Strozzi, the Albizzi were strongly associated with the oligarchic faction in the city. Mentioning the Pitti seems harder to fathom: Buonaccorso Pitti's son Luca was a Medici partisan in 1434. Yet the Pitti, in fact, tended to build very similar portfolios of marriage alliances as the Albizzi and Strozzi during the first 35 years of the 15th century (Padgett 1994). Whether this similarity was translated into a public understanding of their shared placement in social space cannot be known with certainty, but Iacopo's reference to these particular families suggests it might have been. So Iacopo says here, in effect, "These are the kinds of relationships I have sought and found in the past." The point is that self-presentation through patronage letters attempts to establish worthiness discursively while constructing a context of relationships—again discursively—with relevant others by which one's own worthiness or social capital can be inferred and tested. Words get manipulated to present a textured public picture of one's "private" motivations and reliability, and this framing is integral to the achievement of network positions.

FRAMING ACROSS LETTERS: THE DYNAMIC CONSTRUCTION OF RELATIONSHIPS

As the project of self-presentation is inherently designed to change one's structural position for the better, one would like to trace the ways in which this change could be accomplished across a series of letters, rather than simply asserted in one. We can do this for the case of Ormanno di Rinaldo degli Albizzi, son of one of the commune's most prominent politicians, who wrote a series of letters in early 1430 to Averardo de' Medici requesting assistance in securing an office in the naval service of the commune, in particular focusing on the arms he would be allowed to command in the war effort against Lucca. At this time, Averardo was *consolo del mare* in Pisa, a position with considerable influence over such appointments. Ormanno's difficulty was that while he was equal (if not superior) in social status to Averardo, the two found themselves on opposite sides of the factional battles shaping up in the city. Consequently, Ormanno had to begin by making a claim of being served with an office on the basis of honor, not on the basis of *amicisia*. Any claim he might eventually make concerning *amicisia* would have to be actively contextualized.¹⁶

¹⁶ Unfortunately, the current data set of office letters does not contain a group of letters from one supplicant to one patron in which we can witness the gradual construction of an intimate but *hierarchical* relationship, to contrast with the relationship Ormanno constructs with Averardo. The best we can do is a pair of letters from Francesco Arrighi to Averardo of February 4 and February 11, 1430 (MAP II, 175 and 202;

Six of the 10 extant letters Ormanno sent to Averardo were transcribed and are included in figure 2: MAP II, 148, 152, 168, 180, 225, and 226, in chronological order (cases 5, 74, 6, 75, 77, and 78, respectively). Note that the first four of these appear in the "honor" sector, while the fifth appears in the *amicisia* sector, and the last is situated very near the origin. Ormanno's framing shifts during this stream of letters in a way that dynamically constructs his motivations. Custom dictated framing himself as a status inferior, but as we shall see, the rhetoric of the letter often suggests status parity between himself and Averardo. Meanwhile, the intimacy of talk in the letters contrasts oddly with Ormanno's unwillingness to commit himself personalistically to Averardo and his friends. In a sense, then, all four of the cardinal poles of the rhetoric of office seeking identified in figure 1—status parity versus status difference, and strong ties versus weak ties—are activated in this complex correspondence.

Two basic issues arise: (1) How does Ormanno go about framing an image of himself through the narrative, imagery, and cuing in these letters? (2) Why is the ultimate image he constructs a credible one? As to the first of these, the presentation of Ormanno's "true" self follows this basic trajectory: (1) I, Ormanno, assert that I am basically unconcerned with money (salary); (2) more important, I openly disparage the crass politics of *amicisia*, that is, favoritism; I eschew it as a base motivation; (3) I repeatedly adopt the implied high ground of concern for personal and communal honor; but then (4) I rekey (Goffman 1974, chap. 3) my ostentatious rhetoric as mostly playacting and quietly and "self-revealingly" adopt the language and tone of *amicisia*—which I had disparaged—as an acceptable portrait of my motivation, my interests, and my understanding of the way things work. Again, we could not know how extraordinary Ormanno's writing is in the absence of the quantitative analysis of the corpus of letters identified in figure 2, yet we could also not fully understand Ormanno's self-presentation and the nature of its change across letters without observing his rhetorical strategies closely, using concepts borrowed from discourse analysis.

Unfortunately, these letters are too long and involved to be reproduced in full here.¹⁷ Some aspects of them are standard. For example, Ormanno repeatedly expresses certainty that he will obtain Averardo's assistance,

cases 8 and 14 in the honor and service sectors respectively of fig. 2), which are virtually identical in wording concerning a promotion to a new position in Averardo's bank with Andrea de' Bardi, except that Arrighi refers to himself in the conclusion of the second letter as Averardo's servant—as if this subtle framing shift might lend his request more urgency.

¹⁷ Interested readers may consult full transcriptions and translations of these and other office-seeking letters elsewhere (McLean 1996, chap. 6). Transcriptions of some of the Ormanno letters are available in Guasti (1873).

as if to create through the construction of the letter a view of the facts and the alter's character that would of themselves demand that the favor be supplied: "My honor I know I do not have to recommend to you" (case 5); "I have faith in you as in a dear father" (case 5); "I know that you desire that I have honor" (case 74); "I will have entered a new labyrinth if you do not help me—as I have a firm hope you will" (case 74); "I am most certain, wherever you find yourself I cannot have anything other than honor" (case 6); "You being there, [the galley] cannot be other than well armed" (case 75). This is particularly noteworthy near the beginning of the first letter in a place where medieval letter writers would have inserted a *captatio benevolentiae*—a rhetorical gesture designed to flatter and curry the favor of the recipient: Ormanno writes, "You are on the job: I think for all of this the galley will be furnished."¹⁸ Further, Ormanno portrays his relationship with Averardo using the customary imagery of a father-son relationship. The opening salutation of the first two letters and the fourth is "Dearest like a father";¹⁹ in the first letter he writes, "Consider it as if this affair were touching your son, although I don't want so much, but I do have faith in you as in a dear father"; in the second letter we find "in you I trust as in a dear father." This likening of the patron-client role relationship to the father-son role relationship appears repeatedly in Alberti's *Della famiglia* and was ubiquitous throughout Florentine popular culture, perfectly appropriate when the recipient of a letter is considerably older than its sender, as Averardo was considerably older than Ormanno. Finally, Ormanno offers the customary assurances of a willingness to respond to his patron's wishes: "I would be next to you at your every command" (case 5); "I am at your disposal" (case 5); "I would ever be ready for any command of the Dieci" (case 74); "Concluding, again I recommend myself to you" (cases 74, 6, and 75).

But Ormanno is inordinately adamant at the end of his first letter that he is concerned about his personal honor: "You hear the need of my honor, for you initiated it, and likewise you may give it its fulfillment." In so writing, Ormanno links his own sense of honor to that of Averardo, drawing a boundary around the two of them and claiming they are linked by a common motivation leading to a common path of action. This is the third time honor has been mentioned in the first letter, but its salience is

¹⁸ "Voi siete insul fatto. Penso per tutto questo la galea sarà fornita." Note that Ormanno uses the perfect future tense here—not the conditional (*sarebbe fornita*) nor the subjunctive (*sia fornita*), again suggesting certainty of being helped.

¹⁹ However, it is worth noting that 12 other opening salutations in this data set of office letters share this trait of addressing the recipient as "father" or "like a father," including nine others addressed to Averardo (cases 8, 9, 11, 14, 15, 16, 20, 63, and 64).

increased further in the second letter, where Ormanno alludes to his honor fully six times.²⁰ This amplification of the honor frame gives the first two letters a certain urgency: "Perhaps it would be well now more than ever to start, I don't say to conclude things, but to show our intentions and to write to officials, comrades and also the sailors that they take up no other exercises away from here such that when the time comes none of them will be missing." Whereas in the first letter, Ormanno makes the customary claim that he will do "what I think will be pleasing to you and others and to do from now on what goes towards my honor and advantage," in the second letter he exclaims, "I would rather die than turn back at the expense of my honor." Honor now apparently exceeds life itself as a value; how inappropriate it would be now to speak openly of "advantage." The second letter also contains more extravagant pleas for assistance and more interpretive effort from Ormanno concerning the motivations and opinions of various actors.

For certain, Averardo, if this galley be armed with men well-chosen and well-tested, we will have honor; if the opposite, we will face great danger, and you are aware of this. The best thing is always to be prepared. Rinieri Lotti and Banco [Bencivenni], who is here in the galeotta, conspire in not taking care to arm the galley, and with dishonorable words against the last consuls, and against me. It's all wrong; and God knows it, that nothing but honor pulls me here.

Note how starkly Ormanno paints the alternatives here: we will have honor if you follow my advice, shame and humiliation if you do not. And here also, "we" has multiple senses: we the commune will succeed but also we—Ormanno and Averardo—will share in the honor of having done what is best for the commune. The pronoun "we" is used inclusively to link Ormanno's private destiny to that of Averardo and to the publicly prized destiny of the commune (cf. O'Barr 1982). By contrast, Rinieri Lotti and Banco Bencivenni are represented by Ormanno as being outside the

²⁰ Whereas Ormanno repeatedly asserts his obsession with his own honor, his father, Rinaldo, guarded his honor through silences. In his correspondence with the Dieci (a committee delegated directive responsibility for the Lucca war), Rinaldo repeatedly refused their framing of his actions in terms of personal responsibility and the jeopardization of his honor, claiming that he was acting in a way conducive to protecting *their* honor but only his own well-being. When the Dieci claimed to have defended his honor from vicious rumors in the city, Rinaldo's response was not a forceful claim to honor but an implicit denial that his honor could ever have been at issue. In discourse with those outside his own family, Rinaldo does not even countenance a framing of his honor as a point of contention. This difference has obvious methodological relevance for any form of content analysis, for while the presence of keywords and images often signifies active construction of a frame, their absence may likewise be strategic and significant, even though this is not commonly captured effectively through routine statistical and content-analytic methods.

circle of respect for honor: they speak disrespectfully of the previous set of officials and of Ormanno himself.

In the postscript to the second letter, Ormanno continues to distinguish himself from Banco Bencivenni:

I have from Ser Mariotto Bencini, your notary here, that Banco writes a long letter to the consuls, which contains many things at my expense, etc.; and everything he says is the will and opinion of you consuls. I know they are his own; he is a man who sows scandal. I believe he will return here soon; how much more of his bad morals have your friends heard. He is a person who has done badly for himself, and does not take care of the honor of his companion more than his own.

Banco, says Ormanno, misrepresents the opinions of others and treats honor selfishly. His character failings (his "bad morals") are made out to be directly detrimental to the execution of communal policy. Implicitly, Ormanno suggests that he is a different sort of person and more trustworthy than Banco. Ormanno "speaks for another" here (cf. Schifffrin 1993) and relies on the spoken authority of someone whom he can be reasonably sure Averardo will trust, the notary Bencini. Thus, his speaking for another is distinguished from Banco's. Again, he draws an inclusive boundary around himself and Averardo and excludes Bencivenni from the circle of "honorable people," using the renarration of gossip in the city to buttress his claim.²¹

For you were the reason I was elected chief of this blessed galley, I know you desire that I have honor, the which I hope to have by the grace of God, if it be made well-armed, and with experienced men, and not dictated by friendship, as it seems to me Rinieri and Banco wish it. Concluding, again I recommend myself to you. My honor gathers itself now, in arming this galley well from stem to stern.

Note especially in this concluding passage from the end of the second letter how Ormanno asks for a decision in favor of honor, and therefore a decision against friendship (*amicizia*): deserved patronage over partisanship. This juxtaposition and ranking of relevant motivations is actively introduced by the letter's author. Much as in the letter from the notary Iacopo di Antonio di Salvestro to Matteo Strozzi discussed above, the supplicant assertively and contextually establishes the relevant frame for understanding his request: the relevant frame is his (constructed) image of his private motivation—here honor, not partisanship. Patronage is ac-

²¹ Similarly, in his third letter, Ormanno reiterates the need for his allies and supporters to get talking about his appointment, as though talk in itself could bring about a more favorable decision-making climate. In this way, too, he contextualizes his request in terms of the favorable disposition expressed toward it by relevant outside actors (Gumpert 1992).

ceptable if based on the merits of the candidates and their prospects for performing their assigned duties well, and Ormanno argues that a sense of honor, which he possesses in spades, is the prime requisite for the task at hand.

The same message is conveyed slightly differently in the next letter (case 6). Where Ormanno does mention friends here, near the beginning of the letter, they are "your" friends, not "our" friends.²² Toward the end of the letter, "we" (Ormanno and his soldiers, the commune, Ormanno and Averardo) is used inclusively to link Ormanno's personal honor to Averardo's, as it had been in the second letter. A similar use of pronouns occurs in the fourth letter (case 75):

... for you know what must be done with enemies at sea, and especially with armed galleys in the Riviera. I think, by the grace of God, if I go up, I will do honor to whomever has done it to me, and I will make a believer of anyone who speaks badly; for everything proceeds from envy. It pleases me that Messer Rinaldo and you have been together; he wrote of this to me. Some desire his return, and those who wish him ill will desire it most. Even so, there appear to the Dieci for now proofs that he can do some good in this blessed enterprise, in which consists the exaltation of all your friends, and especially considering that it is by the Dieci: what a great reprisal it is to the adversaries, and does them quite a lot of damage.

Ormanno is "inside" with respect to Averardo's honor calculus but outside his circle of "friends." Implicitly, Averardo is entitled to a friendship (read "politically partisan") network—and one clearly existed (Kent 1978)—but Ormanno is not in it. Partisanship and patronage networks, based respectively on favoritism and merit, are constructed as divergent in these letters.²³

Finally, these letters show us something concerning the issue of making and breaking frames. In three letters, Ormanno frames Averardo as a father, as a superior; but halfway through the second letter, he momentarily steps out of this deferential role and says, "Per certo, Averardo"—"for certain, Averardo." This is an address to a peer, not a father; and it cues a plea for honor. So despite the patronal beginning, this is not a letter

²² "It seems there are some interested in seeing me remain on land, even if it's a done deal. When shortly we will be able to speak of it, I will tell you. I am not making an estimate at present, although I would have fifteen days ago, the way things are happening; and it seems to me a thousand years before Messer Rinaldo returns from there, and this is pleasing to all of your friends. May Our Lord conduct things in the best way."

²³ At the end of the fourth letter, Ormanno recommends Messer Batista Duraxino of Genova to Averardo, calling him "our friend" (*nostro amico*), but he goes on to clarify his meaning by saying, "He has always been a friend of the Florentines," underscoring that he uses the notion of friendship here in a military context and "our" in reference to the commune.

sequence about favoritism but about honor and honor's due. Ormanno breaks through the frame set up in the salutation, as if his true self must speak through the decorous forms with which he began. It may seem that this insistence stems from his impetuous and arrogant character, just as he claims in a later letter that all criticism of him stems from envy (*invidia*). But whatever its source, it is as if his inner self is being sincerely represented here. The way this self and its aims are linked to Averardo's evolves across episodes. At the beginning, Ormanno practices a kind of extension technique (Snow et al. 1986); his promise that "I will await whatever they and your companions there will say, and immediately I will obey as much as is assigned to me," communicates the message, "I can conform to your plans and expectations." But later, he seems to say, "Do well by me and you will discover your reward in the honor I will reflect onto you"—that is, "Assuming you are an honorable man, you cannot fail to act in the way I request of you." Stated thus, his approach resembles more Snow's discussion of bridging, where the recipient of a message is invited to bend so as to meet its sender, rather than the sender marketing himself or itself as conforming to the recipient's expectations or desires. Overall, Ormanno fluctuates between these modes of bringing Averardo's interests into congruence with his own, rather as clientage typically oscillates between gestures of deference and defiance (Eisenstadt and Roniger 1984).

Although the first letters set up honor as Ormanno's cardinal motivation, a motivation that implies he will dependably pursue certain kinds of action, this frame is noticeably adjusted in the last two letters here. The opening of the first of these marks a critical moment and bears being reproduced at length.

Until now I have written to you concerning the matter of the galley in one manner, and now I speak in another. It seems to me they intend not to arm the galley for now, which is a great happiness to me. And if my father was not returning, I did not see a way of being able to leave earlier anyway. It seems to me the affair has transpired well, and according to my wishes (*desiderio*). I did not want to refuse, so as not to give pleasure to our enemies (*nostri nemici*), for they would have said that I had done it from cowardice or fear that someone else might be put in my place. If this had been the case, I would not have given so much trouble to [our?] friends (*agli amici*). The fact is, with time, everything comes. I thank you very much for everything you have contrived (*operato*) on my behalf.

"Until now I have written . . . in one manner, and now I speak in another." This sentence betokens a change in frame, a subtly different tactic for speaking and—seemingly—a more honest representation of his motivations. Here it is as if Ormanno modestly reconstructs an interpretation of his earlier presentation of self—a new keying. Earlier, he had taken a stand against threats to (and aspersions on) his honor. Now he shows

himself more patient and willing to let the affair develop as Averardo had contrived it should; and lo and behold, he claims this accords with his "private" wishes. Suddenly, Ormanno offers a "backstage" explanation to the effect that his "onstage" zeal for honor was wrapped up with sustaining appearances, to quell gossip in the city.

Take note: this is the gossip of "our" enemies within the city. Unless his quest were fully honorable (and moderate) in his own eyes, Ormanno now claims, he would not have given so much trouble to his and Averardo's friends. This is post hoc accounting—frame construction on the go—and through it, Ormanno subtly suggests a drift to a more intimate relation with the Medici—whether true (sincere) or not. He says, in effect, "I spoke to you, Averardo, in a way consistent with how I had to act with others in order to maintain my reputation." Like his father before him, Ormanno felt damned if he did and damned if he didn't: by refusing to go unarmed, his enemies would call him a coward; by proceeding rashly, he would stand accused of having too much concern for place and too little sense of civic responsibility. This situation required subtle management. In the last letter, Ormanno simplifies the request: "Concerning the talk one hears here about this galley, I want to beg that you arrange through your office that I be called there. Give unto me the spirit to make Messer Rinaldo happy, and to see that everyone who has exerted himself (*s'è aooperato*) on my behalf will have satisfaction." And only in this last letter does Ormanno use the verb *pregare*, the primary verb in patronage supplication; its absence in earlier letters could hardly fail to communicate to Averardo his expectation of satisfaction and his effective treatment of Averardo as an equal. Only at the end does his writing turn into a properly framed request and come into line with the filial framing used in the earlier letters, by which time, however, he has traversed considerable terrain in laying bear his complex motivations, far surpassing Piero Alberti's simple formula for dealing with different alters. Throughout, Ormanno communicates the context of his action and motivation. Context and motive are not independent of the text, but rather they are constructed and updated through it. It is through discourse as a form of agency that placement in social and cultural space is communicated. Conversely, through the complex representation of the relationship between Averardo and Ormanno, we can sense the tensions and conflicting significances that particular network ties could have.

We are still faced with the challenging task of describing why Ormanno demonstrated this change in tone. Was this elaborate rhetoric (in Wayne Booth's [1983] sense, i.e., a structured, authoritative telling) about honor credible to Averardo? The answer, for any skeptics, is a qualified yes. We do not know if Ormanno went to sea, but (a) we do know from the beginning of MAP II, 168 that Averardo did respond to Ormanno's letters,

(b) we know from other sources (i.e., Ormanno's letters from his own father) that people were talking about his appointment and that Averardo was trying to do something, and (c) we even know that some of the Medici partisans, whose correspondence to Averardo survives, thought that Ormanno's request was worthy of being satisfied.²⁴

In the first place, Ormanno came from a family inordinately attentive to honor, and he likely felt Averardo would respond to an honor-based plea. By contrast, the emerging tensions between the two families meant that Ormanno's request could not be credibly based on a self-portrait claiming partisan *amicizia* with Averardo. The change of tone in the last two letters is not something planned from the outset. Instead it emerges in the course of the interaction in response to external constraints: threats to his honorable reputation from a group of enemies who wanted to find fault with him goaded Ormanno into seemingly rash statements. Inherent in his situation is that Ormanno must orchestrate an image of himself for two audiences simultaneously, offering to only one of them (Averardo) a theatrical aside about his "true" feelings. Once his reputation for honor is salvaged, Ormanno revises the presentation of himself to assure Averardo that they are on the same wavelength. Throughout, the relationship between Ormanno and Averardo is framed through words that texture the image of the letter writer and shape or construct the proper attitude of the reader. Ormanno sought a determinate favor that would allow him to execute his self-image as a man of honor and a patriot of the commune, through a studied activation and structuring of his relationship with an important gatekeeper for favor. In these cases and others, frame assemblage, frame breaking, and narrations of one's experiences and motives

²⁴ Here see letters to Averardo from Mariotto di Francesco Segni (MAP II, 196; case 11 in the service sector of fig. 2), Michelozzo di Bartolomeo Michelozzi (MAP II 201; case 13 in the *amicizia* sector of fig. 2), and Mariotto di Francesco di Giovanni and Riccardo di Niccolò Fagni (MAP II, 221; case 76 in the service sector of fig. 2). Segni's letter interestingly enough tries to piggyback a recommendation for Iacopo di Tommaso di Messer Castellano Frescobaldi onto the endorsement of Ormanno, while Michelozzi's similarly piggybacks a recommendation of his brother onto his understanding of Ormanno's situation. Fagni writes, in part, "at your request, and on account of his *virtù*, we arranged that Ormanno would be elected *padrone* of the galley . . . and having heard this day the Dieci has committed to your office the fact of arming the ship and picking the leader and everything, may we remind you most urgently of your honor and ours, which is that Ormanno command it, and that your office provide everything he needs; for it being otherwise, it would mean a loss of honor for Ormanno and us, and it being ours, so too would it be yours." Thus, Fagni engages in the same play of drawing group boundaries using the concept of honor as did Ormanno earlier, suggesting the group implications of challenges to individual honor. Fagni concludes by reminding Averardo, "We are certain, on account of his *virtù* and his good soul, he will serve the commune most well. And if you were not also certain of this, you would not have spoken as urgently as you did."

comprise techniques that had evolved, in the context of a fluid clientelistic politics of opportunities and favors, to depict motivation and support in interpersonal credibility.

CONCLUSION

This article has aimed, first, to outline a research strategy that interleaves our understanding of culture with our understanding of agency. This means regarding agents as *bricoleurs* putting together building blocks available to them from the cultural background in both formulaic and improvisational ways, in episodes of interaction, with an eye toward building relations with each other. One may say, in other words, that culture happens in the interactions between actors and that it is characterized by an interaction between available cognitive frames and the innovative gestural, written, or conversational agency of these actors. This article has built particularly on the work of Goffman (1974) and his successors for an approach to culture and agency that is sensitive to this reality while also providing us with a list of framing techniques that structure the narrative in particular episodes of discursive interaction.

Rather than simply assuming which frames of meaning commonly show up in strips of activity of a given type, we can use content analysis and multidimensional scaling to provide an image of a cultural space with more than two simple polar conceptual positions and label those positions; and we can inductively generate an image of cultural space that identifies routinely or simply constructed strips of activity and, by contrast, unusual or outlying strips. These techniques would appear useful in discerning empirically (Emirbayer and Goodwin 1994, p. 1442) the structure of a variety of hypothetically imaginable "culture structures" (Kane 1991) and the fabric of a variety of group memberships.²⁵

Further, this article has documented some of the cultural work behind the emergence of social network ties. I demonstrated in table 2 that placement of actors in social networks affected to a noteworthy extent the way

²⁵ Among the settings in which such techniques have already or might in the future be profitably applied are the practices and style of seeking social relationships (*guanxi*) in China, ably discussed by Yang (1994), which may evolve with the shift toward regional economic development and the replacement of party cadres with economically oriented managers in Chinese enterprises; techniques for establishing culturally sensitive business relationships in the global economy; analysis of letters of recommendation or interview processes for the hiring of new personnel, whether in the academy or in business enterprises; and understanding how and why individuals use idiosyncratic verbal signs such as nicknames (cf. Gambetta 1993, chap. 7), slang, or gender stereotypes (cf. Holland and Skinner 1987) to achieve a desired membership within particular social groups.

in which they went about constructing their relations with others. Yet the relationship between structural placement and discursive strategy is not a one-way street. Rather than it being only networks that create an environment in which certain schemata make more sense and are more activated than others, certain schemata must make certain network ties more meaningful or more activated than others. The qualitative analysis undertaken here highlights how social relations are represented and constructed, backgrounded or foregrounded, sometimes reinforced and sometimes transformed, through discourse. The letters effectively tell brief stories about how to construe the relationship between individuals and about individuals' commitment to social groups, using practically and normatively sanctioned framings to accomplish this goal. This storytelling is clearly true of other spheres besides office holding.²⁵ The emergence and solidification of social ties is inherently mediated by representations of the meaning of those ties. Failure to acknowledge the cultural construction of social networks means we see only half the real picture, making it harder for ourselves to understand the noise that inevitably accompanies reading personal loyalties off of network patterns and making the entry of actors into spatially remote networks much harder to explain. Moreover, actors frequently seek opportunities to gain leverage from social groups to which they are already tied by seeking out contact with others across group boundaries. Networks typically provide only a snapshot of this constantly renegotiated reality.

Once we recognize these facts, we are better positioned to understand that entry into network positions is typically more open to competition than we might think, and we are better able to account for success in assuming network positions in terms of the rhetorical and narrative elements utilized in seeking out those positions. The differential expertise of actors in discursive structures has determinate impact on their network structural chances or opportunities. Furthermore, the manner in which actors negotiate their entrance into social ties has determinate consequences for the kind and strength of obligations those ties are able to exert upon them in turn; that is, the choice of frame establishes the terms of one's relationship.

²⁵ For instance, not only does the notion of "marriage" carry with it certain shared understandings, rules, expectations, symbols, and so on, but individual marriages are heavily negotiated ties. See, e.g., the letter of Federigo Sassetti to Niccolò di Iacopo Strozzi of December 4, 1433 (C.S. III, 113:111): "For love of me, take the trouble of being with Neri Bartolini and with Pagolo and Domenico, and exert yourselves together to search out how this girl [Federigo's sister] may be married. Advising you that this is one of the major thoughts I have in this world, I have written of this to Neri and Pagolo and Domenico and if it appears possible to you, together with our other relatives, speak of it with Andrea de' Pazzi." Also see Fabbri (1991).

Civic republicanism was only one of the frames available in Renaissance Florence to represent oneself and to initiate relations with others. Yet whichever frame was used, certain recognizably modern techniques arose—albeit clothed in the unfamiliar rhetoric of honor and magnificence—by which to represent publicly the individual's private motivations. And there is sometimes, though not always, good reason to believe these public representations and posturings corresponded, in the spirit of Adovardo Alberti's remarks at the outset, to very deep constructions of identity. A letter from one Mariotto di Marco to Agnolo di Palla Strozzi on October 20, 1439 (C.S. III, 128:94) alerts us to the constitutive power of relationships. At the end of the letter, Mariotto speaks candidly about his efforts at establishing himself in Florentine society:

And do not be surprised that in my sign-off I adopted such presumption as to ennoble myself a little, in the desire to begin a family name (lineage) in calling myself "dalla palla," for a little while it has come to me to do that which has not been my usual custom, to sign my name in this way, and you know I want to hold onto it, despite the fact that I have seen, and see, all the thousands of knaves each [wanting] to make a lineage, those who do not add one florin of taxes.

This letter, signed "your most humble servant," reminds us that a lineage was a highly valued commodity in Renaissance Florence. It gave one a patina of respectability, a way of establishing an identity in the scheme of communal politics and the commune's history. The family *ricordanza* established one's unique identity in the context of a line of noteworthy historical personages (Branca 1986). All the more surprising then that Mariotto chooses, not a distinctive ancestor, nor a distinctive legacy, nor a distinctive hometown or territory, but his relationship with Agnolo di Palla as the substance of his, so to speak, primary public and personal identity. In a society where a surname is a token of merit, Mariotto chooses a token that can hardly signify personal autonomy. His is an identity crafted, and represented to itself, in relationship.

As Allan Silver has noted, "The conditions of Renaissance political life did not diminish the supreme importance of practical acts as forms of help between friends, but subtilized and complicated the relationship between acts and speech" (1989, p. 288). Words in the Renaissance were a crucial medium for expressing one's intentions and for evaluating the motives of others—but words were not necessarily a reliable mirror of the soul (cf. Neuschel 1989). Further research must investigate empirically how, or whether, more sophisticated techniques of self-presentation emerged as the realization grew of how misleading words could be and as the supply of opportunities for advancement changed. Such research demands not only longitudinal analysis of Florentine patronage networks but also analysis of the rhetorics used outside the political sphere and comparative

work on techniques of self-presentation and the efficacy of these techniques in different historical cases.

REFERENCES

- Alberti, Leon Battista. 1969. *The Family in Renaissance Florence (I libri della famiglia)*, translated by Renee Neu Watkins. Columbia: University of South Carolina Press.
- Alexander, Jeffrey C., and Philip Smith. 1993. "The Discourse of American Civil Society: A New Proposal for Cultural Studies." *Theory and Society* 22:151-207.
- Archer, Margaret S. 1988. *Culture and Agency: The Place of Culture in Social Theory*. Cambridge: Cambridge University Press.
- Baron, Hans. 1966. *The Crisis of the Early Italian Renaissance*. Princeton, N.J.: Princeton University Press.
- Becker, Marvin B. 1968. *Florence in Transition*. Vol. 2 of *Studies in the Rise of the Territorial State*. Baltimore: Johns Hopkins University Press.
- Booth, Wayne C. 1983. *The Rhetoric of Fiction*, 2d ed. Chicago: University of Chicago Press.
- Bourdieu, Pierre. 1977. *Outline of a Theory of Practice*, translated by Richard Nice. Cambridge: Cambridge University Press.
- . 1984. *Distinction: A Social Critique of the Judgment of Taste*, translated by Richard Nice. Cambridge, Mass.: Harvard University Press.
- . 1990. *In Other Words: Essays Towards a Reflexive Sociology*, translated by Matthew Adamson. Cambridge: Polity Press.
- Branca, Vittore, ed. 1986. *Mercanti scrittori: Ricordi nella Firenze tra Medioevo e Rinascimento (Paolo da Certaldo, Giovanni Morelli, Bonaccorso Pitti e Domenico Lenzi, Donato Velluti, Goro Dati, Francesco Datini, Lapo Niccolini, e Bernardo Machiavelli)*. Milano: Rusconi.
- Brown, Penelope, and Stephen C. Levinson. 1987. *Politeness: Some Universals in Language Usage*. Cambridge: Cambridge University Press.
- Brucker, Gene A. 1977. *The Civic World of Early Renaissance Florence*. Princeton, N.J.: Princeton University Press.
- Burt, Ronald S. 1992. *Structural Holes: The Social Structure of Competition*. Cambridge, Mass.: Harvard University Press.
- Cicourel, Aaron V. 1973. *Cognitive Psychology: Language and Meaning in Social Interaction*. New York: Free Press.
- Clifford, James, and George E. Marcus. 1986. *Writing Culture: The Poetics and Politics of Ethnography*. Berkeley and Los Angeles: University of California Press.
- Duranti, Alessandro, and Charles Goodwin, eds. 1992. *Rethinking Context: Language as an Interactive Phenomenon*. Cambridge: Cambridge University Press.
- Eisenstadt, S. N., and L. Roniger. 1980. "Patron-Client Relations as a Model of Structuring Social Exchange." *Comparative Studies in Society and History* 22 (1): 42-77.
- . 1984. *Patrons, Clients, and Friends*. Cambridge: Cambridge University Press.
- Emirbayer, Mustafa, and Jeff Goodwin. 1994. "Network Analysis, Culture, and the Problem of Agency." *American Journal of Sociology* 99 (6): 1411-54.
- Fabbri, Lorenzo. 1991. *Alleanza matrimoniale e patriziato nella Firenze del '400: Studio sulla famiglia Strossi*. Firenze: Leo S. Olschki.
- Faulkner, Robert. 1983. *Music on Demand: Composers and Careers in the Hollywood Film Industry*. New Brunswick, N.J.: Transaction Books.
- Foucault, Michel. 1975. "What Is an Author?" *Partisan Review* 42:603-14.
- Gambetta, Diego. 1993. *The Sicilian Mafia: The Business of Private Protection*. Cambridge, Mass.: Harvard University Press.

- Geertz, Clifford. 1973. *The Interpretation of Cultures*. New York: Basic Books.
- Gergen, Kenneth J., and Mary M. Gergen. 1988. "Narrative and the Self as Relationship." Pp. 17–56 in *Advances in Experimental Social Psychology* 21, edited by Leonard Berkowitz. New York: Academic Press.
- . 1974. *Frame Analysis: An Essay on the Organization of Experience*. New York: Harper Colophon.
- Granovetter, Mark. 1973. "The Strength of Weak Ties." *American Journal of Sociology* 78 (6): 1360–80.
- Grendler, Paul F. 1989. *Schooling in Renaissance Italy*. Baltimore: Johns Hopkins University Press.
- Griffiths, Gordon, James Hankins, and David Thompson, trans. 1987. *The Humanism of Leonardo Bruni: Selected Texts*, vol. 46. Medieval and Renaissance Texts and Studies. Binghamton, N.Y.: Renaissance Society of America.
- Griswold, Wendy. 1987. "A Methodological Framework for the Sociology of Culture." *Sociological Methodology* 17:1–35.
- Guasti, Cesare, ed. 1873. *Commissioni di Rinaldo degli Albizzi*, vol. 3. Firenze: M. Cellini.
- Gumperz, John J. 1992. "Contextualization and Understanding." Pp. 229–52 in *Rethinking Context: Language as an Interactive Phenomenon*, edited by Alessandro Duranti and Charles Goodwin. Cambridge: Cambridge University Press.
- Hanks, William. 1992. "The Indexical Ground of Deictic Reference." Pp. 43–76 in *Rethinking Context: Language as an Interactive Phenomenon*, edited by Alessandro Duranti and Charles Goodwin. Cambridge: Cambridge University Press.
- Herlihy, David, and Christiane Klapisch-Zuber. 1985. *Tuscans and Their Families: A Study of the Florentine Catasto of 1427*. New Haven, Conn.: Yale University Press.
- Holland, Dorothy, and Debra Skinner. 1987. "Prestige and Intimacy: The Cultural Models behind Americans' Talk about Gender Types." Pp. 78–111 in *Cultural Models in Language and Thought*, edited by Dorothy Holland and Naomi Quinn. Cambridge: Cambridge University Press.
- Hood, William. 1990. "Fra Angelico at San Marco: Art and the Liturgy of Cloistered Life." Pp. 108–31 in *Christianity and the Renaissance: Image and Religious Imagination in the Quattrocento*, edited by Timothy Verdon and John Henderson. Syracuse, N.Y.: Syracuse University Press.
- Hunt, Lynn. 1984. *Politics, Culture, and Class in the French Revolution*. Berkeley and Los Angeles: University of California Press.
- Kane, Anne. 1991. "Cultural Analysis in Historical Sociology: The Analytic and Concrete Forms of the Autonomy of Culture." *Sociological Theory* 9:53–69.
- Kent, Dale. 1978. *The Rise of the Medici: Faction in Florence, 1426–1434*. Oxford: Oxford University Press.
- Kent, F. W. 1991. *Bartolommeo Cederni and His Friends: Letters to an Obscure Florentine*. Edited by Gino Corti with F. W. Kent, with an introductory essay by F. W. Kent. Florence: Leo S. Olschki, Istituto Nazionale di Studi sul Rinascimento.
- Kristeller, Paul Oskar. 1983. "Rhetoric in Medieval and Renaissance Culture." Pp. 1–19 in *Renaissance Eloquence: Studies in the Theory and Practice of Renaissance Rhetoric*, edited by James J. Murphy. Berkeley and Los Angeles: University of California Press.
- Laitin, David D., and Aaron Wildavsky. 1988. "Political Culture and Political Preferences." *American Political Science Review* 82 (2): 589–96.
- Lansing, Carol. 1991. *The Florentine Magnates: Lineage and Faction in a Medieval Commune*. Princeton, N.J.: Princeton University Press.
- Linde, Charlotte. 1987. "Explanatory Systems in Oral Life Stories." Pp. 343–66 in *Cultural Models in Language and Thought*, edited by Dorothy Holland and Naomi Quinn. Cambridge: Cambridge University Press.
- McLean, Paul D. 1996. *Patronage and Political Culture: Frames, Networks, and Strat-*

- egies of Self-Presentation in Renaissance Florence*. Ph.D. dissertation. University of Chicago, Department of Political Science.
- Molho, Anthony. 1971. *Florentine Public Finances in the Early Renaissance, 1400–1433*. Cambridge, Mass.: Harvard University Press.
- Muhlhauser, Peter, and Rom Harré. 1990. *Pronouns and People: The Linguistic Construction of Social and Personal Identity*. Oxford: Basil Blackwell.
- Murphy, James J. 1974. *Rhetoric in the Middle Ages: A History of Rhetorical Theory from St. Augustine to the Renaissance*. Berkeley: University of California Press.
- Najemy, John M. 1982. *Corporatism and Consensus in Florentine Electoral Politics, 1280–1400*. Chapel Hill: University of North Carolina Press.
- . 1993. *Between Friends: Discourses of Power and Desire in the Machiavelli-Vettori Letters of 1513–1515*. Princeton, N.J.: Princeton University Press.
- Neuschel, Kristen B. 1989. *Word of Honor: Interpreting Noble Culture in Sixteenth Century France*. Ithaca, N.Y.: Cornell University Press.
- O'Barr, William M. 1982. *Linguistic Evidence: Language, Power, and Strategy in the Courtroom*. New York: Academic Press.
- Padgett, John F. 1994. "Marriage and Elite Structure in Renaissance Florence, 1282–1500." Paper delivered at the Social Science History Association Meetings, Atlanta, Georgia.
- Padgett, John F., and Christopher K. Ansell. 1993. "Robust Action and the Rise of the Medici, 1400–1434." *American Journal of Sociology* 98 (6): 1259–1319.
- Pitt-Rivers, Julian. 1977. *The Fate of Shechem or the Politics of Sex*. Cambridge: Cambridge University Press.
- Sarat, Austin, and William L. F. Felstiner. 1988. "Law and Social Relations: Vocabularies of Motive in Lawyer/Client Interaction." *Law and Society Review* 22 (4): 737–69.
- Schiffrin, Deborah. 1993. "'Speaking for Another' in Sociolinguistic Interviews: Alignments, Identities, and Frames." Pp. 231–55 in *Framing in Discourse*, edited by Deborah Tannen. Oxford: Oxford University Press.
- Schudson, Michael. 1989. "How Culture Works: Perspectives from Media Studies on the Efficacy of Symbols." *Theory and Society* 18 (2): 153–80.
- Sewell, William H., Jr. 1992. "A Theory of Structure: Duality, Agency, and Transformation." *American Journal of Sociology* 91 (1): 1–29.
- Silver, Allan. 1989. "Friendship and Trust as Moral Ideals: An Historical Approach." *Archives européennes de sociologie* 30:274–97.
- Snow, David, E. Burke Rochford, Jr., Steven K. Worden, and Robert D. Benford. 1986. "Frame Alignment Processes, Micromobilization and Movement Participation." *American Sociological Review* 51:464–81.
- Steinmetz, George. 1992. "Reflections on the Role of Social Narratives in Working-Class Formation: Narrative Theory in the Social Sciences." *Social Science History* 16 (3): 489–516.
- Swidler, Ann. 1986. "Culture in Action: Symbols and Strategies." *American Sociological Review* 51 (April): 273–86.
- Tannen, Deborah. 1993. "What's in a Frame? Surface Evidence for Underlying Expectations." Pp. 14–56 in *Framing in Discourse*, edited by Deborah Tannen. Oxford: Oxford University Press.
- Trexler, Richard C. 1980. *Public Life in Renaissance Florence*. New York: Academic Press; Ithaca, N.Y.: Cornell University Press.
- Weissman, Ronald F. E. 1989. "The Importance of Being Ambiguous: Social Relations, Individualism, and Identity in Renaissance Florence." Pp. 269–80 in *Urban Life in the Renaissance*, edited by Susan Zimmerman and Ronald F. E. Weissman. Newark: University of Delaware Press.
- White, Harrison C. 1992. *Identity and Control: A Structural Theory of Social Action*. Princeton, N.J.: Princeton University Press.

- White, Hayden. 1978. *Tropics of Discourse: Essays in Cultural Criticism*. Baltimore: Johns Hopkins University Press.
- Witt, Ronald G. 1988. "Medieval Italian Culture and the Origins of Humanism as a Stylistic Ideal." Pp. 29–70 in *Humanism in Italy*. Vol. 1 of *Renaissance Humanism: Foundations, Forms, and Legacy*, edited by Albert Rabl, Jr. Philadelphia: University of Pennsylvania Press.
- Yang, Mayfair Mei-Hui. 1994. *Gifts, Favors, and Banquets: The Art of Social Relationships in China*. Ithaca, N.Y.: Cornell University Press.

Toward a Role-Theoretic Conception of Embeddedness¹

James D. Montgomery
London School of Economics

Attempting to formalize Granovetter's embeddedness argument, rational choice theorists have viewed social relationships as repeated games. This article argues that role theory would provide a better metatheoretical perspective on embeddedness. A preliminary sketch of role theory suggests a promising theoretical methodology. To illustrate, I construct a repeated-game model in which the players are not individuals but roles (a profit-maximizing "businessperson" and nonstrategic "friend"); the businessperson role acts strategically in light of a metarule that governs intrapersonal role switching.

I. INTRODUCTION

In a widely cited article, Granovetter (1985) argued that economic action is embedded in social relationships. Attempting to formalize this argument, rational choice theorists have viewed social relationships as repeated (prisoner's dilemma) games. However, as discussed in the next section, experimental evidence suggests that cooperation in a prisoner's dilemma follows not from "calculative trust" but from the subjective representation of the situation—the roles evoked in players. Thus, role theory might offer a better metatheoretical perspective on embeddedness. Drawing upon March's (1994) conception of the "logic of appropriateness," I attempt in Section III to sketch the outlines of role theory. The insight that role theory might be viewed as a generalization of game theory suggests a promising theoretical methodology. I illustrate this methodology in Section IV by developing a formal model based loosely on Uzzi's (1996, 1997) ethnographic account of embedded relationships. The formal model resembles the standard finitely repeated prisoner's dilemma, except that the players

¹ This paper has evolved from several preliminary working papers on embeddedness and role theory. For comments and encouragement, I am grateful to Peter Abell, Mark Chaves, Peter Diamond, Tom Fararo, Mark Granovetter, John Padgett, Matt Rabin, Dave Sally, Richard Swedberg, and Brian Uzzi. Direct correspondence to James Montgomery, Interdisciplinary Institute of Management, London School of Economics, London WC2A 2AE, United Kingdom. E-mail: J.Montgomery@lse.ac.uk

are not individuals but roles (a "businessperson" who maximizes profit and a "friend" who feels obligated to cooperate); the businessperson role acts strategically in light of a metarule that governs role switching. Beyond illustrating a methodology that might be applied more widely by role theorists, the model permits a fresh role-theoretic interpretation of several features of embedded relationships described by Uzzi (1996): the distinction between "calculative" and "heuristic" behavior, the need for mutual acquaintances to "prime" relationships, the existence of a "trial phase" characterized by calculative behavior, and cooperation in endgames for nonreputational reasons. I conclude in Section V with general reflections on the nature of trust and the problems raised by rational choice conceptions of social network formation.

II. EMBEDDEDNESS AND REPEATED GAMES

While Granovetter (1985) identified embeddedness as an important *problem* for economic sociology, the *theory* of embeddedness remains vague (Swedberg 1997). Attempting to formalize Granovetter's (1985) argument, rational choice sociologists (e.g., Raub and Weesie 1990) and economists (e.g., Gibbons 1997) have viewed embedded relationships as repeated games. Indeed, because examples of embeddedness typically involve long-term relationships characterized by mutual cooperation and trust in spite of the potential for opportunism (see Granovetter 1985; Coleman 1988), rational choice theorists have tended to equate embeddedness with mutual cooperation in a repeated prisoner's dilemma (PD).² In this view, self-interested actors sustain cooperation through "calculative trust" (Williamson 1993): I trust you because I calculate that your short-run benefit from an opportunistic defection is outweighed by your long-run benefit from continued cooperation.

However, while repeated-game models might seem to offer an elegant account of cooperation in embedded relationships, experimental evidence raises doubts whether cooperation in PDs is based upon calculative trust.³ Game theorists would predict that players will never cooperate in a static (one-shot) PD because defection is a dominant strategy. More generally, recognizing that the last period (endgame) of any finitely repeated PD is essentially a one-shot PD, a simple backward-induction argument implies

² See Gibbons (1992) for the standard game-theoretic analysis of the static and repeated PD. Sociological applications of the repeated PD include Raub and Weesie (1990), Macy (1991), Heckathorn (1996), Lomborg (1996), and Montgomery (1996b).

³ See Dawes and Thaler (1988), Sally (1995), and Rabin (1998) for recent reviews of the experimental literature.

that players should never cooperate in any PD of known finite duration.⁴ Only in the *infinitely* repeated PD—where calculative trust becomes possible—does simple game-theoretic analysis admit the possibility of cooperation.⁵ But in marked contrast to these theoretical predictions, experimental subjects frequently cooperate even in one-shot PDs.⁶ For instance, in one often-cited study, Marwell and Ames (1981) found that subjects in a multiplayer PD (social dilemma) contributed on average between 40% and 60% of their stake to the public good even though a contribution of zero was the dominant strategy for each player. Overall, the experimental evidence strongly rejects self-interest (and, by extension, calculative trust) as a full explanation for behavior in PDs (Dawes and Thaler 1988; Sally 1995; Rabin 1998).

In a metaanalysis of PD experiments, Sally (1995) identified various experimental manipulations (of instructions, group identity, communication, etc.) that alter the cooperation rate. For my present purpose, the manipulation of instructions is of particular interest. Game theorists would typically presume that directives from the experimenter (e.g., “try to help each other” or “maximize your earnings”) are irrelevant; all relevant information is presumed to reside in the payoff matrix. Similarly, neither the labels placed on actions (e.g., “cooperate” and “defect”) nor the label placed on the game as a whole (e.g., “prisoner’s dilemma”) would be expected to influence players’ strategies. However, Sally (1995, p. 78) reports that a directive to cooperate increases cooperation by 34%–40% while a directive to compete decreases cooperation by 20%–33%. Moreover, for repeated PDs, Sally (1995) reports that an affective label (such as “cooperate”) on the dominated choice induces about 20% more cooperation than a neutral label (such as a letter or number). While these results permit various interpretations (see Sally 1995), behavior in PDs clearly depends upon more than the objective specification of the game (given by the payoff matrix). Indeed, these results might be taken as support for the longstanding sociological claim that behavior follows from the subjective “definition of the situation” (Thomas and Znaniecki 1918; Merton 1957).

⁴ The argument runs as follows. Given that both players will defect in the endgame, there can be no long-run benefit from cooperation (to offset the short-run benefit from defection) in the next-to-last period. Thus, both players defect in the next-to-last period. But then there can be no long-run benefit from cooperation in the second-from-last period, and so on. Thus, applying backward induction, the possibility of cooperation “unravels” from the endgame all the way back to first period.

⁵ In a deeper analysis, Kreps et al. (1982) show that cooperation is possible in finitely repeated PDs when there is a small chance that some players follow an “irrational” strategy (such as “tit for tat”).

⁶ Indeed, cooperation rates are typically higher in one-shot PDs than repeated PDs (Dawes and Thaler 1988; Sally 1995).

One particular PD experiment by Samuels and Ross (1993; see also Ross and Ward 1995*a*, 1995*b*) not only demonstrates that subjective representation ("construal") of the situation is an important determinant of behavior but also reveals that the subject's reputation has no predictive power. In this experiment, college dormitory advisors were asked to nominate (from among their advisees) experimental subjects who were thought to be either especially likely or especially unlikely to cooperate in a PD. Subjects were then paired to play a repeated (seven-period) PD. To manipulate subjects' construal of the situation, the experimenter told half of the pairs that they were playing the "Wall Street Game"; the remaining pairs were told that they were playing the "Community Game." But except for this construal manipulation, all pairs faced objectively identical situations.⁷ Given the usual game-theoretic presumption that labels are irrelevant, the experimental results are striking. While only one-third of the subjects in the Wall Street condition cooperated in the first period, more than two-thirds of the subjects in the Community condition did so.⁸ Moreover, while construal clearly influenced behavior, reputation had no predictive power. Within each condition, cooperation rates were virtually identical between those subjects deemed most likely to defect and those deemed most likely to cooperate.⁹

In light of the experimental evidence on cooperation in PDs, economic sociologists might take a fresh look at ethnographic accounts of embedded relationships. Consider, for example, Uzzi's (1996, 1997) description of exchange between manufacturers and contractors in the New York City garment industry. The PD game does reflect (in an admittedly stylized way) the mapping from actions to monetary payoffs: both manufacturers and contractors would profit from joint cooperation but each has an incentive to behave opportunistically. Further, because embedded relationships persist through time, a repeated PD framework may be appropriate. However, the standard analysis of the repeated PD would obscure several in-

⁷ In particular, all pairs faced the same payoff matrix. Moreover, actions were labeled "strategy C" and "strategy D" in both conditions in order to avoid "social desirability" bias.

⁸ Similarly, the overall (seven-period) cooperation rate was 32% among subjects in the Wall Street condition and 68% among subjects in the Community condition. In the endgame, the cooperation rate was 12% among Wall Street subjects and 54% among community subjects.

⁹ Following Samuels and Ross (1993), I have referred to the subject's nomination status ("likely to cooperate" or "likely to defect") as the subject's "reputation." But given that game theorists often equate reputation with a history of actions in a repeated game (see, e.g., Kreps 1990), it may be more appropriate to equate nomination status with the (dormitory advisor's belief about the) subject's type (i.e., payoff function). Thus, the experimental results suggest that individuals do not possess fixed (trans-situational) types—types are induced by situations.

interesting features of these relationships (Uzzi 1996, pp. 679–82). First, exchange partners begin to cooperate only if a mutual friend “primes” the relationship. Second, there is a trial phase before an arm’s-length relationship becomes an embedded relationship. During the trial phase, exchange partners maintain a calculative orientation focused narrowly on economic goals. After the trial phase, exchange partners adopt heuristic decision making and become motivated by friendship as well as profit.¹⁰ Finally, within embedded relationships, individuals cooperate even in endgames. It is important to note that endgame cooperation is not due to reputational concerns (i.e., concerns that an endgame defection against one exchange partner would adversely influence exchange with other partners).

While conceding that none of these observations follows from the standard analysis of the repeated PD, rational choice theorists might insist that extensions of this (very simple and stylized) game could capture Uzzi’s (1996) observations. For instance, to explain why mutual friends are needed to prime new relationships, rational choice theorists might introduce incomplete information about a player’s type (i.e., payoff function).¹¹ Alternatively, to explain why exchange partners gradually become friends and ultimately cooperate in endgames, rational choice theorists might endogenize preferences.¹² However, reflection on experiments like those conducted by Samuels and Ross (1993) might lead economic sociologists away from rational choice theory toward a role-theoretic conception of embeddedness. In contrast to rational choice theory, role theory might presume that individuals do not possess fixed (transsituational) types. Rather, situations evoke various “identities” or “roles” within individuals.¹³ Moreover, to the extent that the rational choice conception of endoge-

¹⁰ As one of Uzzi’s (1997, p. 42) interviewees said, “It is hard to see for an outsider that you become friends with these people—business friends. You trust them and their work. You have an interest in what they’re doing outside of business.” Similarly, another said, “They know that they’re like part of the company. They’re part of the family.”

¹¹ A simple model might assume that some individuals possess a more “trustworthy” type (reflecting a higher discount factor, a higher subjective payoff for cooperation, or a lower subjective payoff for defection) while other individuals possess a less trustworthy type. Although types cannot be directly observed, information about types might be revealed in various ways. For instance, if types are correlated across acquaintances, an individual’s knowledge of a friend’s type may permit the individual to make an inference about a friend-of-a-friend’s type. Compare this with my explanation for why firms hire workers through employee referral (Montgomery 1991).

¹² While arguing for a repeated-game conception of embeddedness, Gibbons (1997) does suggest that some of Granovetter’s (1985) concerns would need to be addressed within the endogenous-preference framework used by Becker and Murphy (1988). See Montgomery (1994) for a sociological application of this framework in which altruism is endogenized.

¹³ See Rabin (1998) for the similar claim that “frames” do not merely induce heuristic errors but actually *determine* preferences.

nous preferences presupposes a coherent and time-consistent self, role theory might develop an alternative conception of "role switching."¹⁴ Thus, while rational choice theorists might view Uzzi's account as an extended PD with endogenous altruism or incomplete information about types, role theorists might see a series of situations in which businessperson or friend roles are evoked.

III. ROLE THEORY

Concerns with rational choice conceptions of embeddedness might lead economic sociologists to consider an alternative role-theoretic perspective. Unfortunately, while the concept of "social role" is ubiquitous in sociological discourse, role theory itself remains fragmented and poorly specified, making problematic any general statement of this perspective.¹⁵ Still, even a cursory description suggests that role theory is a distinct alternative to rational choice theory and might permit more faithful representations of embeddedness. The central premise of role theory is that the individual should be viewed as a collection of social roles. While rational choice theory presumes that individuals are unitary actors with complete and coherent preferences (presumably hardwired into individuals through biological evolution), role theory might presume that roles are evoked by situations and that the content of roles is socially constructed (perhaps changing through a process of social evolution). While rational choice the-

¹⁴ The endogenous-preference framework introduced by Stigler and Becker (1977; see also Becker and Murphy 1988) presumes that individuals possess stable "metapreferences" that permit formation of preferences ("consumption capital") in a time-consistent manner. Experimental evidence that individual behavior is time *inconsistent* (Loewenstein and Elster 1992; Rabin 1988) would seem to undermine the credibility of the Beckerian framework: if individuals make time-inconsistent choices even in settings where preferences can be presumed stable, is it plausible that preferences themselves are formed in a time-consistent manner? O'Donoghue and Rabin (1996) attempt to explain some forms of dynamic inconsistency (e.g., procrastination) by viewing the individual as a sequence of selves; each self plays strategically against future selves. Thus, even within rational choice theory, attempts to explain choice anomalies seem to be leading researchers away from the conception of the individual as a coherent, unitary decision maker.

¹⁵ For reviews of (various sociological versions of) role theory, see Turner (1986, chaps. 20–23), Collins (1988, chap. 7), and Biddle (1986). My (broad) definition of role theory would also encompass recent work on identity (e.g., White 1992). Moreover, role theorists might draw upon related perspectives found in other disciplines including organization theory (see March [1994, chap. 2] on the "logic of appropriateness"), artificial intelligence (see Skvoretz and Fararo [1996] for an AI-inspired conception of rolegrams; see Schank [1990] for a popular account of scripts), formal logic (see Devlin [1997] for a popular account of situation theory), and cognitive psychology (see Rabin's [1998] interpretation of framing effects).

ory presumes that all action derives from utility maximization, role theory might presume that roles contain either rules of behavior or preferences from which those rules might be derived.¹⁶

It is well beyond the scope of the present article to construct an epistemological foundation for role theory. However, the following cursory remarks may help clarify my perspective. Role theory would seem to rest upon two related presumptions: individual behavior presupposes a system of classification, and classification is a necessarily social process.¹⁷ One defense of these presumptions might build upon Wittgenstein's (1958) argument that "private languages"—classification schemes that cannot be communicated to others—are impossibilities.¹⁸ Recognizing that a system of social roles is essentially a system of classifications, Wittgenstein's argument implies that "private" roles are an impossibility—individuals cannot hold a private conception of self that makes no reference to social categories. Rather, individuals can interpret themselves only as a composition of socially constructed roles.¹⁹ Thus, while rational choice theorists would presume that an individual's type is assigned by nature and could be immediately obvious to the individual, role theorists might presume that self-classification ("Who am I?") is a constantly recurring problem of pattern recognition.²⁰

¹⁶ See Padgett and Ansell (1993, p. 1264, n. 11) for a perspective in which preferences exist within (but not across) roles.

¹⁷ It is on these grounds that Douglas (1986) defends the Durkheimian claim that collective entities (such as institutions or societies) can "think." See also Douglas and Hull (1992).

¹⁸ As Padgett (1992, p. 747) explains, individuals have "no access to a third-party view from which to distinguish movements in the world from movements in their own classifications of the world. Objective (i.e., reproducible) knowledge is inherently public." See Bloor (1983) and Collins (1985) for sociology-of-knowledge interpretations of Wittgenstein.

¹⁹ To illustrate, consider an academic engaging in interdisciplinary research. As labeling theory would suggest, the academic may be assigned by others to various social roles (say "economist" or "sociologist") and may subsequently internalize one or more of these roles. But even if the academic is able to overcome this labeling process (perhaps through acute awareness of the possibility of labeling), the academic's self-concept must inevitably be composed only of socially constructed roles. Claims of a private role (neither "economist" nor "sociologist" nor any other social role) would be unintelligible not only to others, but to the academic himself.

²⁰ Rational choice theorists might question why these concerns cannot be addressed within their framework: Why not simply assume that the individual initially possesses incomplete information about her own type and must infer this type from whatever signals are available? (Thus, a worker may only gradually discover whether she possesses high or low ability; a consumer may only gradually discover whether she likes or dislikes a particular product.) In turn, role theorists might dispute whether social categories correspond to "natural kinds" found in human nature. More fundamentally, role theorists might question (at least for many roles) whether there is *any* basis in human nature. For example, on what basis would the academic (see n. 19) infer that

I have been attempting to distinguish role theory from rational choice theory. However, rational choice theory may be more distinct from role theory in rhetoric than in practice. The rhetoric of rational choice theory does, in fact, insist that the individual is a unitary actor. However, except in methodological statements (e.g., Becker 1976), rational choice theorists make no serious attempt to specify the "one big problem" that an individual is purported to solve. Indeed, any specification of this problem (which would require an enumeration of every source of utility, every choice variable, and every constraint faced by the individual) would be overwhelmingly complex. Thus, in practice, rational choice theorists partition the one big problem into more tractable subproblems.²¹ For example, a labor economist might assume that the "worker" faces a simple trade-off between income and leisure. By partitioning the individual's life into subproblems (faced by the "worker," the "consumer," the "parent," etc.), rational choice theorists have (at least in practice) already adopted a role-theoretic perspective in which "types" are dependent upon context.

Before economic sociologists can begin to construct role-theoretic models of embeddedness, they need first a clearer specification of role theory itself. March's (1994, chap. 2) conception of the "logic of appropriateness" may serve as a useful starting point. In this view, a situation evokes one or more roles; these roles contain rules of appropriate behavior. Adopting the label *metarules* for the mapping from situations to roles, and the label *rules* for the mapping from roles to actions, the basic conceptual scheme is illustrated below:



Of course, this initial specification merely creates a demand for further elaboration. What happens if a situation evokes multiple roles (which contain conflicting rules)? What happens if a situation evokes no role (so that an individual does not know how to proceed)? To what extent do individuals *choose* the roles that they play? Relatedly, are metarules a property

he is an "economist" or a "sociologist"? Presumably, neither type is rooted in human biology; both merely represent different interpretations of personal history. In this way, the pattern recognition presumed in role theory seems to resemble more closely Beckerian preference formation (where current preferences are determined by past actions) than imperfect-information models (where types are given by nature but may need to be inferred).

²¹ The decomposition of decision problems leads typically to suboptimization (March 1994, p. 12). Rational choice theorists might defend their practice by arguing that these subproblems are merely useful simplifications that capture the "most important" features of a situation. I would agree, but push the argument further. It is not merely rational choice theorists who cannot completely specify the one big problem: decision makers themselves face exactly the same specification overload.

of individuals (perhaps derived from personal metapreferences) or a property of the social system (and thus determined through a social-level process)? What is the (social-level) process by which roles (and the rules contained within roles) evolve?

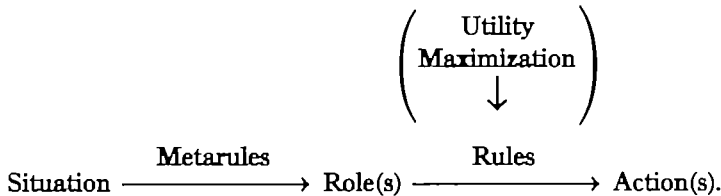
I will not attempt a complete specification of role theory in the present paper but merely note that previous sociological attempts to specify role theory (see Turner 1986, chaps. 20–23) contain a wealth of useful concepts for elaborating the simple scheme illustrated above. Situations where multiple roles are evoked may induce “role conflict”; situations where no roles are evoked may induce “anomie.” Awareness of own past actions creates a “self-concept” that might be merely epiphenomenal (a by-product of self-classification efforts) or might instead provide a foundation for choosing which roles will be played (perhaps constituting a node in a feedback loop running from actions to metarules).²² Finally, while sociologists have made little progress specifying the social-level process by which roles (and the rules within them) evolve (Collins 1988, p. 238; but see Skvoretz and Fararo [1995] for an interesting recent suggestion), I would propose that collective experience of role conflict might lead to a reduction in role content (i.e., the number of rules within roles or the difficulty of following those rules), while collective experience of anomie might lead to an increase in role content.

While a complete specification of role theory must await future research, two elaborations of the simple logic-of-appropriateness scheme will prepare the ground for the model developed in the next section. First, elaborating the process by which situations evoke roles, it may be useful to distinguish an individual’s “focus of attention” from more fundamental ambiguity about the individual’s identity (March 1994, chap. 5). Given that individuals have only limited ability to process information (one aspect of what economists have called “bounded rationality”), an individual might fail (within a particular situation) to consider a relevant role unless this role is brought somehow to her attention. However, even if all relevant roles are brought to the individual’s attention, she may feel that some roles apply only partially. In terms of mathematical formalisms, I am distinguishing between a stochastic process (through which roles are

²² Some role theorists might deny that individuals could choose roles, arguing that individuals possess no “true” or “deep” self (i.e., a self underlying all social roles) that would be capable of making such choices. If so, metarules might change only through social-level (rather than intrapersonal) processes. However, the rhetoric of role theory might suggest that roles are chosen: the ubiquitous claims that individuals “play” or “negotiate” roles would seem to make sense only if there is a deep self underlying (and capable of choosing between) roles. In the limiting case, rational choice theorists might equate the “self-concept” with (Beckerian) metapreferences that provide a stable foundation for (rational) preference formation.

evoked) and a fuzzy set (characterizing the individual's self-assessed "degree of membership" in various social categories).²³

Second, while the logic of appropriateness presumes that roles simply contain rules of behavior, we might generalize this scheme, allowing roles to contain either rules of behavior or utility-maximization problems from which those rules could be derived. In March's (1994) terminology, I am nesting the logic of consequences within the logic of appropriateness. While maintaining that all roles are evoked by situations, I now permit some roles to generate action through rule following (a logic of appropriateness), while other roles might generate action through utility maximization (a logic of consequences). The generalized scheme is illustrated below:



Thus, when constructing models, we might assume that some roles simply contain rules: a citizen should vote, a father should provide for his family, a friend should cooperate. In contrast, other roles might contain preference orderings: a politician maximizes votes, a businessperson maximizes profits.

Although this simple distinction between rule following and utility maximization is convenient for motivating the model below, it appears precarious upon closer inspection and thus requires further comment.²⁴ Rational choice theorists would not deny that rule following is the proximate cause of action, but would merely insist that decision rules are derived from underlying utility-maximization problems. Unless role theorists can develop an empirical methodology that allows them to distinguish between rule following and utility-maximizing behaviors, this distinction merely raises questions about the most useful formal representation of behavior, not the phenomenology of behavior.²⁵ Any well-specified utility-maximiza-

²³ I suspect that the conception of the "individual as fuzzy set of roles" could be used to address many of March's (1994, chap. 5) concerns about ambiguity. See Zadeh (1975) for an early formal treatment of fuzzy sets and fuzzy logic; Kosko (1994) provides a popular account.

²⁴ March (1994) himself refuses to nest one decision-making logic within the other, arguing that it is always possible to offer both logic-of-consequences interpretations of rule following and logic-of-appropriateness interpretations of utility maximization.

²⁵ One might suppose that, on a phenomenological level, most (if not all) behavior derives from rule following. For instance, bookkeeping or religious behavior might be consciously understood as merely the application of a (possibly complex) set of rules. However, these rules themselves may have been "rationalized" either (slowly

tion problem implies decision rules; the question is whether some roles contain rules that either cannot or should not be viewed as the outcome of a utility-maximization problem. Let me offer three arguments suggesting that utility maximization may not always be the best representation of role content.

First, a utility-maximization representation may be trivial. Indeed, rational choice theorists have trouble explaining adherence to "social norms" not because these rules would be difficult to derive from utility maximization, but because the associated maximization problems (entailing merely "psychic" costs and benefits) often seem vacuous. Consequent attempts to specify nontrivial utility-maximization problems may distort our understanding of normative behavior. For instance, while a role theorist might explain voting simply as a rule contained in the "citizen" role, rational choice theorists who have attempted to provide nontrivial explanations have created for themselves the "paradox of voting": Why would any rational person vote when costs exist but expected benefits (from influencing the election outcome) are clearly negligible?²⁶ More generally, the need to construct nontrivial utility-maximization problems have led rational choice theorists to emphasize external sanctions (I follow a norm because others are watching me) over internalization (I follow a norm because of who I am). To the extent that a role has been internalized (i.e., the individual's self-assessed "degree of membership" is positive), the cost of disobeying the norm is found not merely in external sanctions (which are completely irrelevant if the individual knows that she is not being monitored), but in the violation of self-consistency—individuals would no longer "recognize" themselves.²⁷

and haphazardly) through collective practice or (more deliberately) by professional rulemakers (consider the Financial Accounting Standards Board in the case of book-keeping and denominational theologians in the case of religious behavior). Thus, utility maximization might offer a parsimonious representation of role content, even if individuals occupying those roles do not understand the utility-maximization problem from which rules of behavior are derived. More generally, it is important to recognize that *all* metatheoretical frameworks make presumptions. Role theory differs from rational choice theory not by abandoning *as if* assumptions, but by (sometimes) making different ones.

²⁶ Rational choice theorists might protest that, even if the "paradox of voting" has not been resolved, variation in costs (e.g., registration costs) does help explain variation in voter turnout, so that a utility-maximization representation is warranted. Still, I would speculate that variations in voter turnout may ultimately be understood better through "role conflict" (between the "citizen" and other roles) than by positing (stable) costs and benefits within a (unitary) individual.

²⁷ Let me acknowledge that such concerns could be handled (at least partially) within the rational choice framework. Indeed, Montgomery (1994) offers a rational choice interpretation of Liebow's (1967) account in which persistent disobedience to a social norm (provide for your family) eventually causes streetcorner men to abandon a role

Second, experimental findings of "preference reversals" (see Tversky and Thaler [1990] and Rabin [1998] for reviews) might suggest that some sets of decision rules cannot be derived from any simple utility-maximization problem. More generally, cognitive psychologists might suggest that even intendedly rational (i.e., consciously calculative) behavior might be best understood as the application of (sometimes inappropriate) heuristics rather than utility maximization.²⁸

Finally, recognizing that decision rules can be derived only from *well-specified* utility-maximization problems, such representations of role content may create the false impression that an individual's understanding of a situation—the actions available, the mapping from actions to payoffs, the probability that various contingencies will arise—provides a sufficient basis for making a rational choice. In particular, consider the possibility that individuals cannot specify all possible contingencies (states of nature). Rational choice theorists have long recognized the distinction between "risk" (where states are clearly specified and occur with known probability) and more fundamental "uncertainty" (where probabilities are unknown or the states themselves cannot be specified).²⁹ However, because only well-specified utility-maximization problems yield decision rules, rational choice theorists almost always recast uncertainty as risk.³⁰ Accepting that uncertainty pervades human existence, role theorists might represent role contents as rule following in order to emphasize that indi-

(as father). Still, I would speculate that role theory might ultimately provide more complete and less contrived representations of this process.

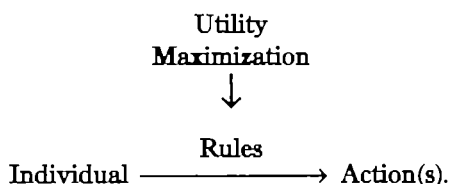
²⁸ In the model below, I assume that the businessperson role contains a profit-maximization problem and derive the behavior of this role using the standard tools of game theory. But note that the "standard tools of game theory" might themselves be understood as merely the collection of heuristics (e.g., backward induction) that game theorists currently deem "rationality." Thus, in an important sense, the rationality attributed to the businessperson is rule following even though the precise rules constituting rationality remain implicit in my analysis of the model. More generally, the findings of cognitive psychologists pose two (open) questions for role theory. First, to what extent are heuristics common across roles? Second, are "preference reversals" generated by cognitive biases within (intendedly rational) roles or by conflicts between roles?

²⁹ See, e.g., Luce and Raiffa (1957, chap. 13); I address this issue in the context of religious choice (Montgomery 1992, 1996a).

³⁰ For instance, Kreps (1996) describes how economic theorists have twisted Williamson's (1975) original conception of bounded rationality (which presumed uncertainty in contracting environments) into a claim about the "verifiability" of performance measures (i.e., whether contracts based upon a particular measure could be legally enforced). Given this twist, "implicit" contracts (based upon nonverifiable measures) as well as "explicit" contracts (based upon verifiable measures) may be represented as well-specified utility-maximization problems. In contrast, role theorists might view an implicit contract as rule-following within an (internalized) role, and an explicit contract as an attempt to induce behavior when a role has not been internalized.

viduals do not possess sufficient information about situations to make rational choices.³¹

One final observation will help motivate the model in the next section. On some (e.g., epistemological) grounds, it may be appropriate to view role theory as an alternative to rational choice theory. However, having elaborated the simple logic-of-appropriateness scheme so that roles may contain either rules or utility-maximization problems, it also becomes possible to view role theory as a generalization of rational choice theory. The basic conceptual scheme in rational choice theory is illustrated below:



Comparing this diagram to the previous one, it appears that the “individual” in rational choice theory occupies the same position as the “role” in role theory. Recognizing this isomorphism, role theorists might interpret rational choice models of individual behavior as models of role behavior. In particular, a game between two individuals becomes a game between two roles.

The recognition that the “players” studied by game theorists could be interpreted as roles rather than individuals would seem to open new avenues of research within game theory (and would also permit role theorists to borrow its well-developed tools). While game theorists have implicitly assumed that each individual is assigned a role that remains fixed throughout the course of a game, role theorists would consider the more general class of games in which role switching is possible. Indeed, to the extent that roles have foresight, the possibility of role switching would seem to complicate strategic interaction. Within dynamic games, each role would need to consider not only the reaction of its current opponent, but also how current actions will affect the roles subsequently adopted by both individuals.³²

³¹ In this way, role theory would follow evolutionary game theory (Weibull 1995). Alternatively, role theorists might retain utility-maximization representations of role content, but explicitly recognize that the subjective representation of a decision environment need not reflect the objective environment.

³² Indeed, in some settings, strategic interaction with own future roles may be more crucial than strategic action against the opponent’s (present or future) roles. Note again the connection to recent work on time inconsistency, which views the individual as a sequence of selves (O’Donoghue and Rabin 1996).

IV. A MODEL

A. Specification

Exploring the idea that role theory represents a generalization of game theory, I now develop a formal model based loosely on Uzzi's (1996, 1997) account of embeddedness. One narrow goal is to derive formally Uzzi's observations, thus providing a role-theoretic interpretation that might be contrasted with standard repeated-game interpretations. But my broader goal is to illustrate a concrete theoretical methodology that role theorists could apply more widely across a range of sociological problems.

To begin, suppose that two individuals are engaged in a finitely repeated PD. Further suppose (in contrast to the standard formulation) that both individuals contain two roles: a "businessperson" seeking to maximize profit and a "friend" who feels obliged to cooperate. Following the distinction made in the previous section, I assume that both roles are evoked in both individuals every period, but that roles may "apply" to individuals to a larger or smaller degree.⁴³ For instance, the friend role may apply only partially if the individuals acknowledge a mutual friend but have had no previous interaction with each other; the friend role might apply to zero degree if the individuals have no mutual acquaintances. Of course, as the game proceeds, the emergent pattern of interaction may itself influence the degree to which a particular role applies. For instance, if one individual defected in the preceding period, the other individual may feel that the friend role no longer applies. Alternatively if one individual continually cooperates, the friend role may become increasingly applicable for the other individual.

Given that both roles are always evoked within each individual, we need to specify how intrapersonal conflicts between these roles would be resolved. (This is the question of role conflict raised in the preceding section.) Recognizing that the present game entails only two roles (which may apply to different degrees) and a binary choice (between cooperate and defect), it may be relatively innocuous to assume that the individual's choice is dictated each period by a dominant role.⁴⁴ Following Uzzi's account, we might assume that the businessperson role is initially dominant,

⁴³ I will not here employ the formalism of a fuzzy set, but that is the underlying idea. The individual is a fuzzy set of roles; the "degree of membership" in each role is between "0" and "1."

⁴⁴ Let me stress that this assumption should not be viewed as a general intrapersonal conflict-resolution mechanism applicable for all games. Indeed, the assumption that a dominant role dictates choice becomes less compelling when individuals contain more than two roles (since a coalition of "weaker" roles might be able to defeat a "stronger" role) or richer action spaces (since conflicting roles might be able to agree upon a "compromise" action).

2

		2	
		C	D
1	C	3, 3	0, 4
	D	4, 0	1, 1

FIG. 1

but that the friend role becomes dominant after some number of periods of cooperation (by self or other or both). Uzzi's account might further suggest that the initial applicability of the friend role (and thus the length of the trial phase) is determined by the strength or pattern of indirect ties between the individuals. Thus, the trial phase would be short if both individuals acknowledge strong ties to a third party; it would be long if there is no chain of social ties connecting the individuals. We might further assume that the businessperson role regains its initial dominance given some number of defections (by self or other or both).

Given that one role dictates choice in each period, we can (at least for the sake of analysis) view the individual as a sequence of roles. To complete the model, I need to specify a particular metarule that governs intra-personal transitions between roles. In the formal analysis, I restrict attention to the following metarule: if an individual never defects and has cooperated at least x times, then the *other* individual is a friend; otherwise, the other individual is a businessperson. Note that this metarule gives each individual the unilateral ability to turn the other individual into a friend. (Of course, as we will see below, the businessperson may find it too costly to do so.) Obviously, many other metarules might be posited; I will consider (informally) the implications of some alternative metarules in the discussion below.

Having developed the model informally, let me now offer a more careful specification. I consider a finitely repeated two-person game. Periods are indexed $1, 2, \dots, T$; the individuals are indexed 1 and 2. In each period, two actions are available to each player: "cooperate" (C) and "defect" (D). The payoffs in each period are given in figure 1.

Each period, each individual is either a businessperson or a friend. A businessperson attempts to maximize the discounted sum of payoffs (over

the course of the entire game) for the individual.³⁵ The businessperson's discount factor is $\beta \in [1/3, 1]$.³⁶ In contrast, a friend follows a fixed strategy: always cooperate.³⁷ Transitions between roles are governed by the following metarule. If individual i has never defected and has cooperated at least x times (where $x < T$), then individual j is a friend; otherwise, individual j is a businessperson.³⁸ (The metarule applies symmetrically, so that $\{i, j\} = \{1, 2\}$ or $\{2, 1\}$.) The businessperson role has foresight and (in particular) is aware of this metarule.³⁹ Note that the model contains three parameters $\{\beta, x, T\}$.

B. Results and Numerical Examples

An analysis of the model is placed in appendix A. The present subsection merely states the major results, offers numerical examples, and provides some intuition.

To summarize the implications of the model, it will be convenient to recast the three parameters of the model as $\{\beta, x, T - x\}$ where β represents the businessperson's discount factor, x represents the length of the trial phase of the game, and $T - x$ represents the length of the friendship phase of the game.⁴⁰ Intuitively, cooperation becomes possible only if the businessperson cares "enough" about future outcomes (i.e., β is close to 1), it doesn't take "too long" to make your opponent into a friend (i.e., x is

³⁵ Thus, even in cases where the individual will switch from businessperson to friend during the course of the game, the businessperson remains concerned with the payoffs that will be received by the friend.

³⁶ Given $\beta \in [0, 1/3]$, one can show that the unique equilibrium sequence of play entails mutual defection throughout the entire game. To simplify the analysis, I thus ignore this case.

³⁷ Having specified the content of the friend role as rule following, let me again acknowledge the precariousness of the distinction between rule following and utility maximization. Instead of specifying the rule "always cooperate," I could just as well specify "subjective payoffs" for the friend such that the friend prefers always to cooperate. Presuming that the businessperson remains concerned only with the "material payoffs" received by the friend, this respecification of the game would not alter the analysis.

³⁸ Note that the game is not interesting if $x \geq T$. In this case, the game never reaches the "friendship" phase; the usual backward-induction argument implies that both individuals will always defect.

³⁹ Given my specification of the friend role, it is irrelevant whether the friend role has foresight.

⁴⁰ Given the maintained metarule, neither individual could become a friend during the trial phase. But note that, depending on actions taken during the trial phase, each individual may or *may not* become a friend at the start of the "friendship phase" of the game.

small), and the friendship phase is "long enough" for players to reap substantial benefits from mutual cooperation (i.e., $T - x$ is large).

To be more precise, analysis of the model reveals that cooperation becomes possible only if one player (say 1) would be willing to cooperate during the trial phase (while 2 defects) in order to turn player 2 from a businessperson into a friend. (In the friendship phase, player 1 then cooperates until the final period so that 2 will continue to cooperate.) If player 1 follows this strategy, she would receive the discounted sum of payoffs (viewed from the beginning of the game) equal to

$$\pi_c = \sum_{t=1}^x \beta^{t-1} 0 + \sum_{t=x+1}^{T-1} \beta^{t-1} 3 + \beta^{T-1} 4.$$

Alternatively, player 1 could simply defect throughout the entire game (while 2 also defects), thus receiving the discounted sum of payoffs:

$$\pi_D = \sum_{t=1}^T \beta^{t-1} 1.$$

Comparing these payoffs, player 1 would be willing to follow the former (make-your-opponent-into-a-friend) strategy only if the difference between these payoffs is nonnegative. Thus, cooperation becomes possible only if

$$\Delta(1) = \pi_c - \pi_D = \sum_{t=x+1}^{T-1} \beta^{t-1} 3 + \beta^{T-1} 4 - \sum_{t=1}^T \beta^{t-1} 1 \geq 0.$$

One can show that this condition holds under the (intuitive) conditions described above: β is sufficiently close to 1, x is sufficiently small, and $T - x$ is sufficiently large.

Assuming that cooperation is possible, analysis of the game is complicated by the existence of multiple equilibria. If we restrict attention to pure strategies, then the equilibrium sequence of play is described in theorem 1. Let $\{a_1, a_2\}$ represent the (ordered) pair of actions taken by individuals 1 and 2 within a particular period. I refer to the final period (T) as the endgame.

THEOREM 1 (equilibrium sequence of play given pure strategies).—If $\Delta(1) < 0$, then 1 and 2 play $\{D, D\}$ for the entire game. If $(x = 2 \text{ and } \beta^{T-2} \geq (1 + \beta)/3\beta)$ or $(x = 1 \text{ and } \beta^{T-1} \geq 1/3)$, then 1 and 2 play $\{C, C\}$ for the entire game. Otherwise, 1 and 2 play $\{C, D\}$ until period x , play $\{C, C\}$ from period $x + 1$ until period T , and play $\{D, C\}$ in the endgame.

To illustrate, consider the following numerical example.

Example 1 (pure strategies).—Suppose $\beta = 0.9$, $x = 3$, and $T = 10$. Given these parameter values, one can verify that $\Delta(1) > 0$. Thus, theorem

Toward a Role-Theoretic Conception of Embeddedness

TABLE 1
SEQUENCE OF PLAY

PLAYER	PERIOD									
	1	2	3	4	5	6	7	8	9	10
<i>i</i>	C	C	C	C	C	C	C	C	C	D
<i>j</i>	D	D	D	C	C	C	C	C	C	C

1 implies the equilibrium sequence of play in table 1. Note that periods 1–3 constitute the trial phase; periods 4–10 constitute the friendship phase. Further note the preceding diagram combines two “mirror image” sequences: $\{i, j\}$ may represent either $\{1, 2\}$ or $\{2, 1\}$.

As example 1 illustrates, friendship is asymmetric when individuals must play pure strategies. Because *i* initially cooperates, *j* becomes a friend in period 4. But because *j* initially defects, *i* remains a businessperson throughout the game. Consequently, *i* defects in the endgame.

Although formal confirmation of this equilibrium requires more careful backward induction (see app. A), the following intuition may be helpful. If *j* defects in period 1, then (the metarule implies that) *i* will remain a businessperson throughout the entire game. Assuming that *j* will defect, *i* is faced with two viable options: either defect immediately or cooperate until the endgame. If *i* defects immediately, then *j* too will remain a businessperson throughout the entire game. In this case (where both *i* and *j* are always businesspersons), the game simply collapses to the standard finitely repeated PD; the usual backward-induction argument implies that both players will defect throughout the entire game. On the other hand, if *i* begins cooperating (and continues to cooperate even though *j* is defecting), then (the metarule implies that) *j* will become a friend at the beginning of the friendship phase (and will thus begin to cooperate). In this case (where *i* is a businessperson and *j* is a friend), *i* would continue to cooperate during the friendship phase until the endgame (when *j*’s reversion to businessperson would no longer matter). Comparing the sums of discounted payoffs associated with these two options, *i* will begin to cooperate since $\Delta(1) \geq 0$. Finally, notice that *j* strictly prefers to defect in period 1 if *i* cooperates. Loosely, *j* gets to “free ride” on *i*’s cooperation in the trial phase, while still receiving the benefits of mutual cooperation in the friendship phase (at least until the endgame).⁴¹

⁴¹ In this way, the game in its entirety resembles a one-shot game of chicken in which $\{C, D\}$ and $\{D, C\}$ are the only pure-strategy Nash equilibria.

Perhaps the sequence of play shown in example 1 does reflect one empirically relevant outcome in exchange relationships. Endgame defections do sometimes occur and might even be accompanied by *i*'s rationalization that she "never really was *j*'s friend" due to *j*'s initial defections. However, because one player always defects in the endgame, such sequences are inconsistent with Uzzi's observation of mutual cooperation in endgames. Attempting to construct a model consistent with Uzzi's observation, we might obviously consider introducing alternative metarules governing role transitions. However, before doing so, we might recall that theorem 1 depends on the restriction that individuals play (only) pure strategies. If we relax this restriction (and presume that individuals begin the game by playing mixed strategies), then there is scope for joint cooperation in both the trial phase and endgame.⁴²

THEOREM 2 (equilibrium sequence of play given mixed strategies).—*If $\Delta(1) < 0$, then 1 and 2 play $\{D, D\}$ for the entire game. If $[x = 2 \text{ and } \beta^{T-x} \geq (1 + \beta)/3\beta]$ or $(x = 1 \text{ and } \beta^{T-x} \geq 1/3)$, then 1 and 2 play $\{C, C\}$ for the entire game. Otherwise, 1 and 2 play a mixed strategy in period 1; both cooperate with probability $p_1 \in (0, 1)$. If the outcome $\{C, C\}$ has occurred in all previous periods, then 1 and 2 continue to play mixed strategies until the mixed-strategy equilibrium no longer exists (i.e., until period $x - 1$ if $\beta^{T-x} \geq (1 + \beta)/3\beta$, until period x if $(1 + \beta)/3\beta > \beta^{T-x} \geq 1/3$, or until period $x + 1$ otherwise). Individuals 1 and 2 then play $\{C, C\}$ for the remainder of the game. However, if the use of mixed strategies yields the outcome $\{C, D\}$ before period $x + 1$, then $\{C, D\}$ is repeated until period $x + 1$, $\{C, C\}$ is played from period $x + 1$ until period T , and $\{D, C\}$ is played in the endgame. Similarly, if the use of mixed strategies yields the outcome $\{D, C\}$ before period $x + 1$, then $\{D, C\}$ is repeated until period $x + 1$, $\{C, C\}$ is played from period $x + 1$ until period T , and $\{C, D\}$ is played in the endgame. Finally, if the use of mixed strategies yields the outcome $\{D, D\}$ before period $x + 1$, then $\{D, D\}$ is repeated for the remainder of the game.*

To illustrate, consider the following example (which assumes the same parameter values used in example 1).

Example 2 (mixed strategies).—Suppose $\beta = 0.9$, $x = 3$, and $T = 10$. Given these parameter values, one can verify that $(1 + \beta)/3\beta > \beta^{T-x} > 1/3$. One may further verify that $\Delta(1) > 0$. Thus, theorem 2 implies that 1 and 2 will begin by playing mixed strategies in period 1 and will do so again in period 2 if $\{C, C\}$ is played in period 1. I derive these mixed

⁴² Note again the resemblance of the present game to the one-shot game of chicken. The game of chicken generates not only the pure-strategy equilibria $\{C, D\}$ and $\{D, C\}$ but also a mixed-strategy equilibrium. If players mix strategies, then the symmetric outcomes $\{C, C\}$ and $\{D, D\}$ will occur with positive probability.

Toward a Role-Theoretic Conception of Embeddedness

TABLE 2
SEQUENCES OF PLAY GIVEN MIXED STRATEGIES

Sequence	Probability
DDDDDDDDDD	
DDDDDDDDDD	$(1 - p_1)^2 = .0714$
CCCCCCCCCD	
DDDDCCCCC	$2p_1(1 - p_1) = .3917$
CDDDDDDDD	
CDDDDDDDD	$p_1^2(1 - p_1)^2 = .0034$
CCCCCCCCCD	
CDDCCCCC	$2p_1^2p_2(1 - p_1) = .0792$
CCCCCCCCC	
CCCCCCCCC	$p_1^2p_1^2 = .4542$

strategies in appendix B, showing that $p_2 = .9198$ and $p_1 = .7327$.⁴³ Given these probabilities, it is possible to enumerate every sequence that might be played, along with the corresponding probability that it is played (table 2). Note that I have again simplified by combining "mirror image" sequences.

In comparing examples 1 and 2, the (unique) equilibrium sequence of play in example 1 is merely one possible outcome (occurring with probability .392) in example 2. More generally, asymmetric friendships and end-game defections remain possible when players choose mixed strategies. However, symmetric friendships and mutual cooperation in endgames also become possible. Indeed, the final sequence—mutual cooperation through the entire game—occurs with probability .454.

C. Discussion of the Model

Armed with these derivations, I can now give a role-theoretic interpretation of Uzzi's (1996) observations. The distinction between "calculative" and "heuristic" behavior arises because the businessperson role generates action through utility maximization while the friend role generates action through rule following.⁴⁴ The "trial period" observed at the beginning of a new relationship reflects the initial dominance of the calculative busi-

⁴³ Note that $p_1 < p_2$. More generally, p_i will be increasing as long as mixed strategies exist. That is, conditional upon mutual cooperation in the current period, each player is more likely to cooperate in the next period.

⁴⁴ Admittedly, this distinction remains precarious in the present model; recall n. 37.

nessperson role. While cooperation by one individual could eventually change the other individual from a businessperson into a friend, cooperation (and thus friendship) arises in equilibrium only if the trial phase is not too long. Thus, to explain why mutual acquaintances are needed to prime new relationships, one might argue that mutual acquaintances make the friend role more applicable at the start of an exchange relationship, thus shortening the trial phase. Finally, if we consider the equilibrium sequence of play generated by mixed strategies, mutual cooperation may arise in endgames as a consequence of symmetric friendship.

Of course, the possibility remains that another model (which assumed a different metarule) could support a more compelling interpretation of Uzzi's (1996) findings. In particular, one might search for alternative metarules that will guarantee mutual cooperation in the endgame. Although I will not provide formal analyses, it may be instructive to consider briefly the implications of two alternative metarules. First, recall the metarule presumed in the present model:

METARULE.—*If i has never defected and has cooperated at least x times, then j is a friend; otherwise, j is a businessperson.*

While this metarule gives each player unilateral control over the other player's type, we might consider metarules that give players unilateral control over their own type. For instance,

METARULE A.—*If i has never defected and has cooperated at least x times, then i is a friend; otherwise, i is a businessperson.*

Because metarule A implies that i (once she became a friend) would never retaliate against a defection from j , this metarule would encourage opportunistic behavior by j (which would, in turn, discourage i from becoming a friend in the first place). The equilibrium sequence of play would thus entail mutual defection throughout the entire game. Reflection on the implications of metarule A might lead us to consider metarules in which each individual's type is contingent upon the actions of *both* players. For instance,

METARULE B.—*If both i and j have never defected and if both have cooperated at least x times, then i is a friend; otherwise, i is a businessperson.*

Given metarule B, both players have a strong incentive to cooperate during the trial phase: while mutual cooperation implies that both will become friends (and thus cooperate for the remainder of the game), unilateral defection would imply that both individuals remain businesspersons (and thus defect for the remainder of the game). Indeed, mutual cooperation arises in this case for much the same reason it arises in the standard *infinitely repeated* PD; metarule B takes the place of trigger strategies. However, while this metarule does imply cooperation in the endgame, one might note that the distinction between the trial phase and friendship

phase no longer possesses any empirical content. (Analogously, there is no trial phase in the standard infinitely repeated PD.) More formally, mutual cooperation remains an equilibrium sequence of play even as x becomes very large. Thus, while metarule B does imply cooperation in endgames, we can no longer explain why mutual acquaintances (which shorten the trial phase) would be needed to prime relationships.

Having offered an interpretation of Uzzi's (1996) observations, let me now recall my larger goal: to provide a concrete theoretical methodology that might be used by role theorists to construct formal models. As discussed in Section III, role theory might be viewed as a generalization of game theory in which players are roles rather than individuals. Thus, while role theorists might borrow many of the tools of standard game theory, they would specify not a unique "type" for each individual, but rather an entire repertoire of roles that each individual might play. Consequently, role theorists would need to specify the intrapersonal processes by which roles are evoked and (potential) conflicts between roles are resolved. My model has illustrated one possible method that could be employed more widely by role theorists: assume that roles are evoked sequentially and that transitions between roles are governed by an exogenous metarule. Admittedly, this methodology may not be rich enough to encompass many (potentially important) aspects of role theory indicated in Section III.⁴⁵ Still, while more foundational work on role theory continues, the concrete methodology illustrated here might help researchers at least begin to explore the implications of a role-theoretic perspective.

V. SUMMARY AND DISCUSSION

Attempting to formalize Granovetter's (1985) argument, rational choice theorists have viewed embedded relationships as repeated games. But reflection on this conception of embeddedness, especially in light of the experimental evidence on cooperation in PDs, might lead economic sociologists to explore alternative metatheoretical approaches. Role theory, by relaxing the presumption that the individual is a unitary actor, would permit more flexible (and perhaps less contrived) representations of embeddedness. Although the present paper has provided only a partial sketch of role theory, the insight that role theory might be viewed as a generalization of game theory has suggested a promising theoretical methodology. To illustrate, I have constructed a formal model that resembles the stan-

⁴⁵ For instance, my assumption of a dominant role in each period may obscure interesting ideas about multivocality (Padgett and Ansell 1993). Moreover, even my cursory discussion of alternative metarules might suggest that the key task facing role theorists is to specify the (social-level?) process by which metarules evolve.

dard finitely repeated PD except that the players are roles (a profit-maximizing businessperson or a nonstrategic friend) rather than individuals; the businessperson role chooses actions in light of a metarule that governs role switching. Beyond illustrating a methodology that might be applied more widely by role theorists, the model has permitted a fresh role-theoretic interpretation of several features of the embedded relationships described by Uzzi (1996).

I close with some general reflections. Responding to criticisms that rational choice theorists have misunderstood the (noncalculative) nature of trust, Williamson (1993) has distinguished between "calculative" and "personal" trust.⁴⁶ However, rational choice practice continues to presuppose that calculative trust (or, following Williamson, simply "calculation") is sufficient for understanding economic behavior.⁴⁷ More fundamentally, the metatheoretical presumption that all action derives from utility maximization makes it impossible even to define noncalculative trust within the rational choice framework. In contrast, because the content of social roles might be represented as either rule following or utility maximization, role theory enables more general definitions. Thus, while game theorists have implicitly defined (calculative) trust as an expectation of behavior based upon common knowledge of payoffs and rationality, role theorists might define trust as an expectation of behavior based upon common knowledge of metarules (governing role transitions) and rule following or utility maximization (within roles). Note that both personal and calculative trust emerge as special cases. I may trust a friend because I know the metarules that turn individuals into friends (e.g., a past history of mutual cooperation) and the rules that friends follow (e.g., cooperate); I may trust a businessperson because I know the metarules that turn individuals into businesspersons (e.g., work experience or business school) and the preferences held by businesspersons (e.g., profit maximization).

Perhaps rational choice theorists have been quick to dismiss personal trust as a basis for economic action (see Williamson 1993) largely because

⁴⁶ Indeed, Williamson (1993) regards "calculative trust" an oxymoron, and ultimately agrees with March and Olsen (1989, p. 27) that "the core idea of trust is that it is not based on an expectation of its justification. When trust is justified by expectations of positive reciprocal consequences, it is simply another version of economic exchange." Note that March and Olsen (1989, p. 27) go on to elaborate the role-theoretic view that trust is "a rule of appropriate behavior. It is sustained by socialization into the structure of rules, and rarely considered as a deliberate willful action."

⁴⁷ Indeed, Williamson (1993) explicitly claims that personal trust is rare. Similarly, Gibbons (1997) distinguishes "assurance" from "trust" but uses only the former concept to analyze cooperation within and between firms.

they have been too optimistic about calculative trust. That is, given multiple equilibria in a repeated game, rational choice theorists typically focus attention on equilibria of the repeated game that represent Pareto improvements over the equilibrium of the one-shot game. For instance, when applying the infinitely repeated PD to substantive problems, rational choice theorists typically focus on the equilibrium in which mutual cooperation is forever repeated, ignoring the continuum of alternatives (including the equilibrium in which mutual defection is forever repeated). Given that role theory would enable a clearer representation of personal trust, role theorists might also be more pessimistic about calculative trust. In particular, given an infinitely repeated PD between individuals who are irrevocably calculative, self-interested businesspersons, role theorists might presume that the cooperation could never get off the ground. Rather, cooperation would emerge only if these businesspersons could somehow transform themselves into (altruistic) friends.

This last comment prompts more general reflection on the nature of social ties. In particular, should social ties be represented as choice variables? In the model developed above, the friend role is a fixed strategy that always cooperates. Recognizing the material benefits of mutual cooperation, two businesspeople might wish to jump immediately into a symmetric friendship. But because these role transitions are not "legitimate" (i.e., do not occur through the application of a socially entrenched meta-rule), these transitions are not binding on the businesspeople; friendship (and its benefits) cannot be so directly willed. More generally, role theorists might emphasize that individuals adopt roles (such as friend) through a process of self-classification into these (socially constructed) categories. Thus, while rational choice models might assume that social ties are the direct product of choice (Boorman 1975; Rotemberg 1994; Jackson and Wolinsky 1996; Montgomery 1996b), role theorists might assume that social ties are the indirect product of a longer sequence of actions taken by the self and others (as illustrated in the model above).

Further reflection on the metarules that govern role transitions suggests that the instrumental formation of social ties leads to more fundamental paradoxes. Granovetter (1988, p. 211; following Blau 1964) argues that "a perception by others that one's interest in them is mainly a matter of 'investment' will make this investment less likely to pay off; we are all on the lookout for those who only want to use us." If individuals foresee the material benefits of friendship, but refuse to become friends with anyone who would benefit materially from their friendship, then how is friendship possible? In this light, the possibility that businesspersons can (even indirectly) choose to become friends appears as problematic as Pascal's (1966) suggestion that an individual can choose to form a (materially

beneficial) belief in God by going through the motions of religious observance (Elster 1979, 1983; Padgett 1986; Montgomery 1992, 1996a).⁴⁸ A formal statement and resolution of such choice-theoretic paradoxes would seem to be an important direction for future role-theoretic research.⁴⁹

To conclude, let me suggest that a deeper understanding of embeddedness awaits a clearer and more complete specification of role theory. Rational choice theorists have arrived at an understanding of calculative trust because they possess a well-specified theoretical methodology—game theory—that allows representation and analysis of rational choice explanations. As I have illustrated, role theorists might begin to understand embeddedness by generalizing this methodology. But I suspect that the full force of Granovetter's (1985) argument will be felt only after a complete specification of role theory—the (intrapersonal) process by which role conflict and anomie are resolved, the (intrapersonal or social) process by which metarules evolve, the (social) process by roles (and role contents) evolve—permits more faithful representation and rigorous analysis of role-theoretic explanations.

APPENDIX A

Analysis of the Model

CONDITIONS.—*At the beginning of each period, one of the following conditions will hold:*

- (C, C) both individuals have never defected.
- (C, D) player 1 has never defected; 2 has defected at least once.
- (D, C) player 1 has defected at least once; 2 has never defected.
- (D, D) both individuals have defected at least once.

Note that period $\pi + 1$ is the period in which each individual may experience a transition from businessperson to friend. Recall that period T is

⁴⁸ Padgett and Ansell (1993, p. 1263) argue that the solution to such paradoxes lies in "multivocality—the fact that single actions can be interpreted coherently from multiple perspectives simultaneously, the fact that single actions can be moves in many games at once, and the fact that public and private motivations cannot be parsed." Thus, even if the activation of one role (e.g., "businessperson") seems to preclude a second role (e.g., "friend"), there may exist a third role (perhaps "friend of a friend"?) through which the second role could be activated.

⁴⁹ Blau's (1964) discussion of "social approval" suggests that an individual's preference for social approval is defined over other's *beliefs* rather than others' *actions*. Consequently, psychological games (which generalize standard games by allowing payoffs to be defined over not only actions but also beliefs, beliefs about beliefs, etc.) may be a useful formalism. See Geanakoplos, Pearce, and Stacchetti (1989) for the seminal paper and Rabin (1993) for an application. See also the related discussion in Rabin (1998).

the final period (i.e., the endgame). Let $\{a_1, a_2\}$ represent the (ordered) pair of actions taken by 1 and 2.

LEMMA 1 (period $x + 1$).—*If the following condition holds at the beginning of period $x + 1$, then the following sequence of play occurs in equilibrium:*

- (C, C) the individuals play {C, C} for the remainder of the game.
- (C, D) they play {C, C} until period T , then play {D, C} in the endgame.
- (D, C) they play {C, C} until period T , then play {C, D} in the endgame.
- (D, D) they play {D, D} for the remainder of the game.

Proof of lemma 1

(C, C).—Because 1 has cooperated x times, 2 becomes a friend. Conversely, because 2 has cooperated x times, 1 becomes a friend. Because friends always cooperate, both individuals will cooperate for the remainder of the game.

(C, D).—Because 1 has cooperated x times, 2 becomes a friend. Because 2 has defected in at least one previous period, 1 remains a businessperson. If 1 ever defects against 2, 2 will revert to businessperson; backward induction implies that both 1 and 2 would then defect for the remainder of the game. Individual 1 will wait until the endgame to defect if $3 + 4\beta \geq 4 + \beta$. This condition simplifies to the maintained assumption $\beta \geq 1/3$.

(D, C).—This is the converse of (C, D).

(D, D).—Because both individuals have defected, both remain businesspersons. Backward induction implies that both will defect for the remainder of the game. QED

We can now proceed by applying backward induction. Consider period x .

LEMMA 2 (period x).—*If the following condition holds at the beginning of period x , then the following equilibria arise in the period- x stage game:*

- (C, C) If $\beta^{T-x} \geq 1/3$, then the unique equilibrium is {C, C}. Otherwise, there are two pure-strategy equilibria, {C, D} and {D, C}, and a mixed-strategy equilibrium in which both individuals cooperate with probability $p_x \in (0, 1)$.
- (C, D) The unique equilibrium is {C, D}.
- (D, C) The unique equilibrium is {D, C}.
- (D, D) The unique equilibrium is {D, D}.

To proceed further using backward induction, we need to select the equilibrium that will arise if the condition (C, C) holds at the beginning of period x and $\beta^{T-x} < 1/3$. I will first consider the case in which an asym-

metric equilibrium arises in period x . (Without loss of generality, I will assume that player 2 defects while 1 cooperates.) I will then return to the case in which 1 and 2 play mixed strategies.

Equilibrium selection rule 1.—If both $\{C, D\}$ and $\{D, C\}$ are equilibria of the period- t stage game, assume that $\{C, D\}$ is played.

The following definition will also prove useful in subsequent lemmas:

$$\begin{aligned}\Delta(t) &= \sum_{j=x+1}^{T-1} \beta^{T-j} 3 + \beta^{T-t} 4 - \sum_{j=t}^T \beta^{T-j} 1 && \text{if } x+1 < T; \\ &= \beta^{T-t} 4 - \sum_{j=t}^T \beta^{T-j} 1 && \text{if } x+1 = T.\end{aligned}$$

For lemmas 3 and 5, it is useful to note that $\Delta(x-1) \geq 0$ implies $\beta^{T-x} \geq (1+\beta)/3\beta$. For lemmas 4 and 6 and theorems 1 and 2, it is useful to note that $\Delta(t)$ is increasing in t : $\Delta(x-1) > \Delta(x-2) > \dots > \Delta(1)$.

LEMMA 3 (period $x-1$, pure strategies).—*Assume equilibrium selection rule 1. If the following condition holds at the beginning of period $x-1$, then the following equilibria arise in the period $x-1$ stage game:*

- (C, C) If $\beta^{T-x} \geq (1+\beta)/3\beta$, then the unique equilibrium is $\{C, C\}$. If $\beta^{T-x} < (1+\beta)/3\beta$ and $\Delta(x-1) \geq 0$, then the equilibrium $\{C, D\}$ is played. Otherwise, the unique equilibrium is $\{D, D\}$.
- (C, D) If $\Delta(x-1) \geq 0$, then the unique equilibrium is $\{C, D\}$. Otherwise, the unique equilibrium is $\{D, D\}$.
- (D, C) If $\Delta(x-1) \geq 0$, then the unique equilibrium is $\{D, C\}$. Otherwise, the unique equilibrium is $\{D, D\}$.
- (D, D) The unique equilibrium is $\{D, D\}$.

Continuing with backward induction, we may now state a general result for all periods $t \leq x-2$.

LEMMA 4 (period $t \leq x-2$, pure strategies).—*Assume equilibrium selection rule 1. If the following condition holds at the beginning of period t , then the following equilibrium will arise in the period- t stage game:*

- (C, C) If $\Delta(t) \geq 0$, then the equilibrium $\{C, D\}$ is played. Otherwise, the unique equilibrium is $\{D, D\}$.
- (C, D) If $\Delta(t) \geq 0$, then the unique equilibrium is $\{C, D\}$. Otherwise, the unique equilibrium is $\{D, D\}$.
- (D, C) If $\Delta(t) \geq 0$, then the unique equilibrium is $\{D, C\}$. Otherwise, the unique equilibrium is $\{D, D\}$.
- (D, D) The unique equilibrium is $\{D, D\}$.

Proofs of lemmas 2–4

Detailed proofs of these lemmas are tedious, but available from the author upon request. Here, I merely describe the recursive structure of these proofs. Given the condition that holds at the beginning of period x , the pair of actions taken by players in period x will determine the condition that holds at the beginning of period $x + 1$. Lemma 1 then reveals equilibrium play (and payoffs) from period $x + 1$ onward. Thus, given lemma 1, we may derive, for each of the four conditions that could hold at the beginning of period x , the equilibrium (or equilibria) of the period- x stage game. These results are stated in lemma 2. Next, given the condition that holds at the beginning of period $x - 1$, the pair of actions taken by players in period $x - 1$ will determine the condition that holds at the beginning of period x . Lemmas 2 and 1 then reveal equilibrium play (and payoffs) from period x onward. Thus, given lemmas 2 and 1 (and applying equilibrium selection rule 1 whenever we need to select among multiple equilibria), we may derive, for each of the four conditions that could hold at the beginning of period $x - 1$, the equilibrium (or equilibria) of the period $x - 1$ stage game. These results are stated in lemma 3. Finally, lemma 4 builds upon lemmas 3, 2, and 1 in a similar way. After proving that lemma 4 holds for period $t = x - 2$, it is straightforward to demonstrate (by induction) that this lemma holds for all periods $t < x - 2$. QED

Recognizing that condition (C, C) holds in period 1, theorem 1 follows immediately from lemmas 1–4 and equilibrium selection rule 1. (See text for statement of theorem 1.) Returning to lemma 2 (period x), we might adopt an alternative equilibrium selection rule, and then once again proceed using backward induction.

Equilibrium selection rule 2.—If both {C, D} and {D, C} are (asymmetric) pure-strategy equilibria of the period- t stage game, then assume that the (symmetric) mixed-strategy equilibrium is played.

LEMMA 5 (period $x - 1$, mixed strategies).—*Assume equilibrium selection rule 2. If the following condition holds at the beginning of period $x - 1$, then the following equilibria arise in the period $x - 1$ stage game:*

- (C, C) If $\beta^{T-x} \geq (1 + \beta)/3\beta$, then the unique equilibrium is {C, C}. If $\beta^{T-x} < (1 + \beta)/3\beta$ and $\Delta(x - 1) \geq 0$, then 1 and 2 play mixed strategies, cooperating with probability $p_{x-1} \in (0, 1)$. Otherwise, the unique equilibrium is {D, D}.
- (C, D) If $\Delta(x - 1) \geq 0$, then the unique equilibrium is {C, D}. Otherwise, the unique equilibrium is {D, D}.
- (D, C) If $\Delta(x - 1) \geq 0$, then the unique equilibrium is {D, C}. Otherwise, the unique equilibrium is {D, D}.
- (D, D) The unique equilibrium is {D, D}.

For period $t \leq x - 2$, we can prove the mixed-strategy analog of lemma 4.

LEMMA 6 (period $t \leq x - 2$, mixed strategies).—*Assume equilibrium selection rule 2. If the following condition holds at the beginning of period $t \leq x - 2$, then the following equilibria will arise in the period- t stage game:*

- (C, C) If $\Delta(t) \geq 0$, then 1 and 2 play mixed strategies, cooperating with probability $p_t \in (0, 1)$. Otherwise, the unique equilibrium is {D, D}.
- (C, D) If $\Delta(t) \geq 0$, then the unique equilibrium is {C, D}. Otherwise, the unique equilibrium is {D, D}.
- (D, C) If $\Delta(t) \geq 0$, then the unique equilibrium is {D, C}. Otherwise, the unique equilibrium is {D, D}.
- (D, D) The unique equilibrium is {D, D}.

Proofs of lemmas 5 and 6

Detailed proofs are tedious but available from the author upon request. The proof of lemma 5 follows closely the proof of lemma 3, except that equilibrium selection rule 2 is applied in place of equilibrium selection rule 1. Similarly, the proof of lemma 6 follows closely the proof of lemma 4, except that lemma 5 replaces lemma 3 and that equilibrium selection rule 2 replaces equilibrium selection rule 1. To prove both lemmas 5 and 6, it is helpful to specify payoffs recursively. That is, letting $\pi_i^{t,s}$ represent the discounted sum of payoffs received by player 1 in period t when $\{a_1, a_2\}$ is played, it is helpful to specify period $t - 1$ payoffs as

$$\pi_{t-1}^{CC} = 3 + p_t \beta \pi_t^{DC} + (1 - p_t) \beta \pi_t^{DD};$$

$$\pi_{t-1}^{CD} = 0 + \beta \pi_t^{CD};$$

$$\pi_{t-1}^{DC} = 4 + \beta \pi_t^{DC};$$

$$\pi_{t-1}^{DD} = 1 + \beta \pi_t^{DD}.$$

Recognizing that condition (C, C) holds in period 1, theorem 2 follows immediately from lemmas 1, 2, 5, and 6 and equilibrium selection rule 2. (See text for statement of theorem 2.)

APPENDIX B

Derivation of Mixed Strategies in Example 2

To begin, consider period 2. Given each pair of actions $\{a, a\}$ that might be taken in period 2, lemmas 1 and 2 (see app. A) determine the equilibrium sequence of play from period 3 onward, and thus determine the sum of discounted profits that accrue to player 1 (fig. B1). Substituting for $\beta = 0.9$, we obtain the payoffs shown in figure B2. Thus, 1 prefers to defect

	C	D
C	$\sum_{s=2}^{10} \beta^{s-2} 3$	$0 + \beta^0 + \sum_{s=4}^9 \beta^{s-2} 3 + \beta^8 4$
D	$4 + \beta^4 + \sum_{s=4}^9 \beta^{s-2} 3 + \beta^8 0$	$\sum_{s=2}^{10} \beta^{s-2} 1$

FIG. B1

	C	D
C	18.3774	13.1079
D	18.9859	6.1258

FIG. B2

	C	D
C	$3 + \beta(17.9548)$	$\sum_{s=1}^3 \beta^{s-1} 0 + \sum_{s=4}^9 \beta^{s-1} 3 + \beta^9 4$
D	$\sum_{s=1}^3 \beta^{s-1} 4 + \sum_{s=4}^9 \beta^{s-1} 3 + \beta^9 0$	$\sum_{s=1}^{10} \beta^{s-1} 1$

FIG. B3

	C	D
C	19.1593	11.7970
D	21.0874	6.5132

FIG. B4

if 2 cooperates, and 1 prefers to cooperate if 2 defects. Given that the payoff matrices for 1 and 2 are symmetric, we thus find two pure-strategy equilibria: {C, D} and {D, C}. To compute the mixed-strategy equilibrium, assume that 2 plays C with probability p_2 (and plays D with probability $1 - p_2$). Individual 1 is willing to mix only if

$$18.3774p_2 + 13.1079(1 - p_2) = 18.9859p_2 + 6.1258(1 - p_2),$$

which implies $p_2 = .9198$. Thus, if the use of mixed strategies yielded {C, C} in period 1, both 1 and 2 will choose C in period 2 with probability $p_2 = .9198$. Note that the expected value of playing this mixed strategy is

$$18.3774(.9198) + 13.1079(.0802) = 17.9548.$$

Now consider period 1. Applying the relevant lemmas (1, 2, and 5), one may determine the equilibrium sequence of play from period 2 onward (given the period-1 outcome) and thus construct the payoff matrix for player 1 in period 1 as shown in figure B3. Substituting for $\beta = 0.9$, we obtain the matrix shown in figure B4. Again, given the symmetry of the payoff matrices for 1 and 2, the game yields the two pure-strategy equilibria {C, D} and {D, C}. There is also a mixed-strategy equilibrium in which both 1 and 2 choose C with probability p_1 where

$$19.1593p_1 + 11.7970(1 - p_1) = 21.0874p_1 + 6.5132(1 - p_1),$$

which implies $p_1 = .7327$. Thus, 1 and 2 begin the game in period 1 by choosing to cooperate with probability .7327.

REFERENCES

- Becker, Gary S. 1976. *The Economic Approach to Human Behavior*. Chicago: University of Chicago Press.
- Becker, Gary S., and Kevin M. Murphy. 1988. "A Theory of Rational Addiction." *Journal of Political Economy* 96:675-700.

Toward a Role-Theoretic Conception of Embeddedness

- Biddle, Bruce J. 1986. "Recent Developments in Role Theory." *Annual Review of Sociology* 12:67-92.
- Blau, Peter. 1964. *Exchange and Power in Social Life*. New York: Wiley.
- Bloor, David. 1983. *Wittgenstein: A Social Theory of Knowledge*. London: Macmillan.
- Boorman, Scott A. 1975. "A Combinatorial Optimization Model for Transmission of Job Information through Contact Networks." *Bell Journal of Economics* 51:574-82.
- Coleman, James S. 1988. "Social Capital in the Creation of Human Capital." *American Journal of Sociology* 94:S95-S120.
- Collins, H. M. 1985. *Changing Order: Replication and Induction in Scientific Perspectives*. Chicago: University of Chicago Press.
- Collins, Randall. 1988. *Theoretical Sociology*. San Diego: Harcourt Brace Jovanovic.
- Dawes, Robyn M., and Richard H. Thaler. 1988. "Anomalies: Cooperation." *Journal of Economic Perspectives* 2:187-97.
- Devlin, Keith. 1997. *Goodbye, Descartes: The End of Logic and the Search for a New Cosmology of the Mind*. New York: Wiley.
- Douglas, Mary. 1986. *How Institutions Think*. Syracuse, N.Y.: Syracuse University Press.
- Douglas, Mary, and David Hull. 1992. *How Classification Works: Nelson Goodman among the Social Sciences*. Edinburgh: Edinburgh University Press.
- Elster, Jon. 1979. *Ulysses and the Sirens: Studies in Rationality and Irrationality*. Cambridge: Cambridge University Press.
- . 1983. *Sour Grapes: Studies in the Subversion of Rationality*. Cambridge: Cambridge University Press.
- Geanakoplos, John, David Pearce, and Ennio Stacchetti. 1989. "Psychological Games and Sequential Rationality." *Games and Economic Behavior* 1:60-79.
- Gibbons, Robert. 1992. *Game Theory for Applied Economists*. Princeton, N.J.: Princeton University Press.
- . 1998. "Notes on Two Horse Races: Hobbes and Coase Meet Repeated Games." In *Trust in Organizations*, edited by K. Cook. New York: Russell Sage Foundation, in press.
- Granovetter, Mark. 1985. "Economic Action and Social Structure: The Problem of Embeddedness." *American Journal of Sociology* 91:481-510.
- . 1988. "The Sociological and Economic Approaches to Labor Market Analysis: A Social Structural View." Pp. 187-216 in *Industries, Firms, and Jobs*, edited by George Farkas and Paula England. New York: Aldine de Gruyter.
- Heckathorn, Douglas D. 1996. "The Dynamics and Dilemmas of Collective Action." *American Sociological Review* 61:250-77.
- Jackson, Matthew O., and Asher Wolinsky. 1996. "A Strategic Model of Social and Economic Networks." *Journal of Economic Theory* 71:44-74.
- Kosko, Bart. 1994. *Fuzzy Thinking*. London: Flamingo.
- Kreps, David M. 1990. "Corporate Culture and Economic Theory." Pp. 90-143 in *Perspectives on Positive Political Economy*, edited by J. Alt and K. Shepsle. New York: Cambridge University Press.
- . 1996. "Markets and Hierarchies and (Mathematical) Economic Theory." *Industrial and Corporate Change* 5:561-95.
- Kreps, David M., Paul Milgrom, John Roberts, and Robert Wilson. 1982. "Rational Cooperation in the Finitely Repeated Prisoners' Dilemma." *Journal of Economic Theory* 27:245-52.
- Liebow, Elliot. 1967. *Tally's Corner*. Boston: Little, Brown.
- Loewenstein, George, and Jon Elster, eds. 1992. *Choice over Time*. New York: Russell Sage Foundation.
- Lomborg, Bjørn. 1996. "Nucleus and Shield: The Evolution of Social Structure in the Iterated Prisoner's Dilemma." *American Sociological Review* 61:278-307.

- Luce, R. Duncan, and Howard Raiffa. 1957. *Games and Decisions*. New York: Wiley.
- Macy, Michael W. 1991. "Learning to Cooperate: Stochastic and Tacit Collusion in Social Exchange." *American Journal of Sociology* 97:808-43.
- March, James G. 1994. *A Primer on Decision Making*. New York: Free Press.
- March, James G., and Johan P. Olsen. 1989. *Rediscovering Institutions*. New York: Free Press.
- Marwell, Gerald, and Ruth Ames. 1981. "Economists Free Ride, Does Anyone Else?" *Journal of Public Economics* 15:295-310.
- Merton, Robert K. 1957. *Social Theory and Social Structure*. New York: Free Press.
- Montgomery, James D. 1991. "Social Networks and Labor Market Outcomes: Toward an Economic Analysis." *American Economic Review* 81:1408-18.
- . 1992. "Pascal's Wager and the Limits of Rational Choice: A Comment on Durkin and Greeley." *Rationality and Society* 4:117-22.
- . 1994. "Revisiting Tally's Corner. Mainstream Norms, Cognitive Dissonance, and Underclass Behavior." *Rationality and Society* 6:462-88.
- . 1996a. "Contemplations on the Economic Approach to Religious Behavior." *American Economic Review* 86:443-47.
- . 1996b. "The Structure of Social Exchange Networks: A Game-Theoretic Reformulation of Blau's Model." *Sociological Methodology* 26:193-225.
- O'Donoghue, Ted, and Matthew Rabin. In press. "Doing It Now or Later." *American Economic Review*.
- Padgett, John F. 1986. "Rationally Inaccessible Rationality." Review of Jon Elster's *Sour Grapes*. *Contemporary Sociology* 15:26-28.
- . 1992. "Learning from (and about) March." *Contemporary Sociology* 21:744-49.
- Padgett, John F., and Christopher K. Ansell. 1993. "Robust Action and the Rise of the Medici, 1400-1434." *American Journal of Sociology* 98:1259-1319.
- Pascal, Blaise. 1966. *Pensées*. New York: Penguin.
- Rabin, Matthew. 1993. "Incorporating Fairness into Game Theory and Economics." *American Economic Review* 83:1281-1302.
- . 1998. "Psychology and Economics." *Journal of Economic Literature* 36:11-46.
- Raub, Werner, and Jeroen Weesie. 1990. "Reputation and Efficiency in Social Interactions: An Example of Network Effects." *American Journal of Sociology* 96:626-54.
- Ross, Lee, and Andrew Ward. 1995a. "Naïve Realism in Everyday Life: Implications for Social Conflict and Misunderstanding." Pp. 103-35 in *Values and Knowledge*, edited by T. Brown, E. Reed, and E. Turiel. Hillsdale, N.J.: Erlbaum.
- . 1995b. "Psychological Barriers to Conflict Resolution." 27:255-304 in *Advances in Experimental Social Psychology*, edited by Mark P. Zanna. San Diego, Calif.: Academic Press.
- Rotemberg, Julio. 1994. "Human Relations in the Workplace." *Journal of Political Economy* 102:684-717.
- Sally, David. 1995. "Conversation and Cooperation in Social Dilemmas: A Meta-Analysis of Experiments from 1958 to 1992." *Rationality and Society* 7:58-92.
- Samuels, Steven M., and Lee Ross. 1993. "Reputations versus Labels: The Power of Situational Effects in the Prisoner's Dilemma Game." Unpublished manuscript. Stanford University.
- Schank, Roger C. 1990. *Tell Me a Story: A New Look at Real and Artificial Memory*. New York: Scribner's.
- Skvoretz, John, and Thomas J. Fararo. 1995. "The Evolution of Systems of Social Interaction." *Current Perspectives in Social Theory* 15:275-99.
- . 1996. "Generating Symbolic Interaction: Production System Models." *Sociological Methods and Research* 25:60-102.

Toward a Role-Theoretic Conception of Embeddedness

- Stigler, George J., and Gary S. Becker. 1977. "De Gustibus Non Est Disputandum." *American Economic Review* 67:76–90.
- Swedberg, Richard. 1997. "New Economic Sociology: What Has Been Accomplished, What Is Ahead?" *Acta Sociologica* 40:161–82.
- Thomas, W. I., and F. Znaniecki. 1918. *The Polish Peasant in Europe and America: Monograph of an Immigrant Group*. Boston: Badger.
- Turner, Jonathan H. 1986. *The Structure of Sociological Theory*. Chicago: University of Chicago Press.
- Tversky, Amos, and Richard H. Thaler. 1990. "Preference Reversals." *Journal of Economic Perspectives* 4:201–11.
- Uzzi, Brian. 1996. "The Sources and Consequences of Embeddedness for the Economic Performance of Organizations: The Network Effect." *American Sociological Review* 61:674–98.
- . 1997. "Social Structure and Competition in Interfirm Networks: The Paradox of Embeddedness." *Administrative Science Quarterly* 42:35–67.
- Weibull, Jorgen. 1995. *Evolutionary Game Theory*. Cambridge, Mass.: MIT Press.
- White, Harrison C. 1992. *Identity and Control: A Structural Theory of Social Action*. Princeton, N.J.: Princeton University Press.
- Williamson, Oliver E. 1975. *Markets and Hierarchies*. Free Press: New York.
- . 1993. "Calculativeness, Trust, and Economic Organization." *Journal of Law and Economics* 36:453–86.
- Wittgenstein, Ludwig. 1958. *Philosophical Investigations*. New York: Macmillan.
- Zadeh, Lotfi A. 1975. "Calculus of Fuzzy Restrictions." Pp. 1–39 in *Fuzzy Sets and Their Applications to Cognitive and Decision Processes*, edited by Lotfi A. Zadeh, King-Sun Fu, Kokichi Tanaka, and Masamichi Shimura. New York: Academic Press.

Rethinking Age Dependence in Organizational Mortality: Logical Formalizations¹

Michael T. Hannan
Stanford University

This article explores the use of logical formalization to clarify an area of research characterized by conflicting claims and divergent empirical findings. The substantive focus concerns the relation between organization age and the hazard of mortality. The literature contains claims that the hazard (a) falls with age (a "liability of newness"), (b) rises initially and then falls with age (a "liability of adolescence"), (c) rises with age ("liabilities of senescence and obsolescence"). The formalizations reported cast the relevant theoretical arguments as propositions involving five concepts: endowment, imprinting, inertia, capability, and position. It shows that each of the theoretical stories can be derived as implications of particular assumptions within two broad formalizations. This analysis clarifies the mechanisms at work in each theoretical account and provides guidance for empirical research designed to discriminate among the competing theories.

I. INTRODUCTION

This article uses the tools of rational reconstruction and logical formalization to reexamine issues of aging and organizational mortality. This reexamination is prompted by increasing divergence in empirical findings about what was regarded as an established fact: the hazard of organizational mortality is higher for young organizations than for old ones (see below). Previous efforts to explain this pattern and the efforts to explain recent divergent findings have yielded a set of partly congruent and partly conflicting theoretical stories. The analysis presented here tries to make

¹ This research was supported by a grant from the Alfred P. Sloan Foundation and by the Stanford Graduate School of Business Faculty Trust. I am grateful to Glenn Carroll, Joel Podolny, Ezra Zuckerman, and the members of the Stanford Workshop on Organizational Ecology for helpful suggestions and to Gábor Péli, László Pólos, and Jaap Kamps for extensive advice on the formalization. Of course, these colleagues are not responsible for the errors that remain. Direct correspondence to Michael T. Hannan, Graduate School of Business, Stanford University, Stanford, California 94305-5015.

sense out of the resulting welter of claims. It reconstructs the arguments in terms of a set of common theoretical constructs, translates them into the language of first-order logic (FOL), and derives their implications.

One goal of this effort is methodological and aims to show that the strategy of logical formalization can play a valuable role in theory building in a context of a set of conflicting theories. The recent successful uses of this methodology in sociology have focused on evaluating the consistency, soundness, and other properties of a particular theory as stated in single text. So, for instance, Péli et al. (1994), Péli (1997), Péli and Masuch (1997) have formalized various theory fragments in organizational ecology, and Pólos and Kamps (1997) formalized a portion of Thompson's (1967) theory. These efforts have clarified the logic of the theories, sharpened their assumptions, and suggested previously unrecognized implications and better understandings of the ideas. This article applies the methodology to an area of inquiry that now lacks canonical theory. Instead of examining the implications of a single argument, it seeks to clarify how the various alternative stories "work," what must be assumed to yield the opposing claims. Formalization has two potential advantages in this kind of setting. First, by identifying the assumptions that really do the work in the different stories, such formalization can provide guidance to empirical research, directing attention to relationships that need special scrutiny. Second, it can serve as the basis for building new theories that have the alternative stories as clearly defined special cases.

A second goal is substantive. This involves developing more fully a distinction that lies at the heart of much sociological analysis of organizations but has not received much theoretical attention: how capability-based advantages and positional advantages combine to shape organizational life chances. In particular, this formalization builds on a distinction between fragile and robust positions and specifies the difference in terms of the degree to which advantage accruing to position generalizes across states of the environment. This kind of distinction yields implications about organizational mortality processes that resonate with the results of recent empirical research.

A. Substantive Background

Two images coexist uneasily in standard sociological analyses of organizational mortality. One regards the structural features of organizations as mappings from the environment encountered at founding (Stinchcombe 1965), and it depicts organizations as locked into their initial forms. Structural inertia dominates this image: organizations have very limited capacities to reshape their core structures (their forms) as quickly as the environment changes (Hannan and Freeman 1977). Efforts at change in core

features diminish life chances, at least in the short run (Hannan and Freeman 1984). Therefore, when environments are variable and uncertain, extensive change in the distribution of organizational forms occurs by selection processes operating on organizational populations.

The second view concentrates on the life-history dynamics of organizations. It portrays young organizations as particularly vulnerable to environmental selection—they face a liability of newness (Stinchcombe 1965). Aging conveys advantages, such as improved capabilities and more secure structural positions, that tend to insulate older organizations from damage due to environmental turbulence. Hence the life chances of organizations improve with aging.

The two images do not necessarily conflict. For instance, large shifts in the environment might increase mortality rates for all organizations in a population and also expose the youngest organizations to the most intense force of mortality. This said, it must be admitted that the two theoretical images do not fit comfortably. Structural inertia in the face of environmental change ought to erode fitness. If new organizations can incorporate current understandings, best practices, and state-of-the-art technology in their core structures and if old organizations have the core structures that reflect a bygone era, why do old organizations have better life chances?

The two images became paired within sociological theories of organization because theorists regarded the liability of newness as an empirical fact that must be accommodated. Carroll (1983), Freeman, Carroll, and Hannan (1983), and much subsequent research showed that the liability of newness held for an impressive range of populations of firms and other kinds of organizations.¹ Thus early organizational ecology research assumed negative age dependence as a baseline model for specifying ecological hypotheses. Hannan and Freeman (1984) derived the liability of newness from an evolutionary theory of structural inertia and organizational change.¹

Subsequent research showed that the relationship between age and mortality rates sometimes takes a more complicated form. Mortality rates sometimes begin low, then rise after a brief interval before settling into

¹ Both papers did, however, note that the evidence was not unequivocal, that the apparent liability of newness might reflect the operation of age-invariant unobserved heterogeneity.

¹ The logical formalization of the Hannan-Freeman argument by Péli et al. (1994) also derived the liability of newness hypothesis from their rationally reconstructed versions of Hannan and Freeman's assumptions and theorems. However, Péli, Pólos, and Hannan (1997) provide two other logical formalizations that better reflect the original argument. In these new formalizations, the liability of newness does not necessarily follow as an implication of the formalized Hannan and Freeman theory of structural inertia.

a pattern of long-term decline with age. This pattern has been dubbed "a liability of adolescence" (Brüderl and Schüssler 1990; Fichman and Levinthal 1991; Brüderl, Preisendörfer, and Ziegler 1996).

As long as the belief in the empirical validity of the claim that mortality rates decline with aging (perhaps beyond some initial interval) remained unshaken, there was little impetus to scrutinize the seeming incongruence between the stories based on inertia and those based on age-related improvements in capabilities and positions. However, the empirical picture has begun to change. It has become recognized that the research establishing the generality of negative age dependence (beyond an initial period) failed to control for the effects of variations in size over organizational lifetimes. This neglect matters because size variations might account for the observed low mortality rates of old organizations: organizations tend to grow as they age, and mortality rates decline with size. For instance, Barron, West, and Hannan (1994) reported results that turned the estimated effect of age on mortality rates from (significant) negative to (significant) positive in one organizational population. And, they proposed that the important liabilities involve processes of senescence and obsolescence.

Evidence on the effects of controlling for size in studies of age dependence has been mounting rapidly. Hannan et al. (1998) identified 18 studies that report the results of analyses that relate organizational mortality rates to age and size and that also update size repeatedly over the life spans of all (or nearly all) organizations in a population. Five studies find *monotonic negative* effects of age, even after controlling for age-varying organizational size. Four studies find a variation on the pattern of non-monotonic age dependence: an *inverted-U shape pattern* of age dependence after controlling for age-varying size. The remaining nine studies find *monotonic positive* age dependence once age variations in organizational size have been taken into account. What do these 18 studies imply? If one combines studies reporting monotonic negative effects of age on mortality rates and those showing inverted-U shaped effects of age on mortality rates (because both imply that the mortality rates of old organizations fall below those of younger ones), the result is a draw: nine studies find positive age dependence and nine studies find (long-run) negative age dependence.⁴

All 18 studies share the assumption that age and size have *proportional effects* on mortality rates. Hannan et al. (1998) argue that such an assumption is not warranted; they find that effects of size and age on mortality

⁴ The exact numbers given here ought not to be taken too seriously because the 18 studies differ considerably in the quality of data and estimation; see Hannan et al. (1998).

rates in populations of automobile manufacturers in four countries are indeed nonproportional. They suggest that the divergence among the earlier studies might be due to the nonproportionality in the effect combined with heterogeneity in the size distributions among the populations studied. These findings hint that the existing theoretical formulations need revision.

This divergence in empirical findings demands rethinking of the issue of age dependence in organizational mortality. The available theoretical accounts (entailing liabilities of newness, adolescence, senescence, and obsolescence) cannot all be true generally, because they disagree about the form of age dependence in processes of organizational mortality. Can these theories be reconciled and unified? Can conditions be specified that tell when one theory applies and others do not? At the present state of development of the theories, we lack clear answers to these questions. The informal nature of the argumentation makes it hard to discern the assumptions invoked in each account. Moreover, it is not even clear that each account is internally consistent, that the conclusions follow from the premises. Continued theoretical progress likely requires that the informal arguments be made more precise and formal and that systematic tools be used to evaluate the theoretical claims.

This article, which builds upon an empirical effort by Hannan et al. (1998), tries to sort out the relevant theories and use the strategy of logical formalization to clarify some basic differences among them. It seeks to make explicit the tacit assumptions that inform the various theories about age dependence so that we might better understand the interplay between arguments about age dependence and general sociological theories of organizations. Following Péli et al. (1994), Péli (1997), and Pólos and Kamps (1997), the formalizations of the competing theories use the tools of first-order logic (FOL). The first step in the analysis, sometimes called rational reconstruction, tries to represent the essence of each argument in stark and simple terms, as a set of assumptions and proposed theorems. The second step checks whether the proposed theorems follow logically from the assumptions.

The stage of rational reconstruction of the informal theories in formal language would ideally work from a single set of axioms thought to be true in all circumstances. However, the relevant sociological arguments do not reflect agreement about such first principles. This means that one cannot begin with the same premises in constructing each theoretical argument without radically changing the spirit of at least one of the contending arguments. Instead of imposing unification at the outset, this analysis seeks to formalize each account on its own terms. In doing so, I try to find constructive interpretations. That is, when several alternative ren-

derings of the verbal argument are available, I adopt the ones that can allow the conclusions to follow from the premises (Péli 1997). The goal of the rational reconstruction is formalizing each theory as a logically consistent system stated in terms that are sufficiently similar to guide subsequent efforts at unification.

II. DECOMPOSING THE ARGUMENTS

The first step in formalizing the theories is to survey the theoretical terrain. The relevant theoretical arguments can be built from propositions concerning five concepts: endowment, imprinting, inertia, capability, and position (Hannan et al. 1998).

A. Endowments

Organizations differ in the quantities and qualities of their initial resources. Some get endowed with extensive financial and social capital, because their founders have great wealth, status, or political influence or because the social conditions of founding are favorable (resources might be abundant and few corporate actors might be competing for them). Others find themselves severely disadvantaged at founding, with these conditions reversed. Endowments bear on the issues under consideration in two ways. First, developing capabilities is costly, and extensive endowments permit greater investment in capability. Carroll and Hannan (1989) argue that organizations founded in periods of intense competition for resources face a liability of scarcity. They cannot invest in developing capabilities, and the failure to make such investments at the outset has irreversible negative consequences for life chances. Second, endowments can affect mortality rates directly. A well-endowed organization can maintain its structures and members even if it cannot continually mobilize resources from the environment. Endowments matter most in the first months and years of operation. Endowments depreciate unless replenished by continuing positive flows of resources from the environment. Until an endowment has been depleted, the organization's risk of mortality is low. Once the endowment gets exhausted, the risk of mortality jumps. The implications for age dependence in mortality rates are clear. At a given level of endowment, the mortality rate remains low during the period of depletion, and it jumps afterward. That is, mortality rates increase with age. The usual formulation of this idea (Brüderl et al. 1996) regards the level of endowment as an unobservable random variable, which makes the length of the period of exhaustion also an unobservable random variable. If the distribution of endowments is (roughly) continuous, one will observe that

the hazard of mortality rises smoothly from zero to a peak (the point of exhaustion for the best-endowed organizations).

B. Imprinting and Structural Inertia

Imprinting refers to a process in which events occurring at certain key developmental stages have persisting—possibly lifelong—consequences. The idea that firms and other kinds of organizations tend to be imprinted by their founding conditions comes from Stinchcombe's (1965) insight that social and economic structures have their maximal impact on new organizations. In attempting to accumulate financial and human capital, entrepreneurs expose their designs to intense scrutiny. Proposals get tested against taken-for-granted assumptions about structural forms and employment relations. Because conventional wisdoms and taken-for-granted assumptions change over historical time (as new forms flourish and others wane), the tests imposed on proto-organization also change. Consequently, the kinds of organizations that emerge reflect the social structure of the founding period. Imprinting requires an initial mapping of an environmental condition onto the nascent corporate actor. The imprinted characteristics must be inert (or at least possess a fair degree of hysteresis). Otherwise, subsequent modifications of the structure will erode the association of founding conditions and those features.

What does imprinting imply about age dependence in mortality rates? If imprinting occurs, then founders build organizations that fit historically specific environments. If core features of organizations get set by early decisions and actions and resist change afterward, then environmental change will erode the fit between organizations and environments. Barron et al. (1994) assume that the distance of an organization's current environment from its founding-period environment varies directly with its age. Then, the quality of the match declines monotonically with age, and age dependence is positive. In other words, the joint action of imprinting, inertia, and environmental change create a liability of obsolescence.

C. Capability

An organizational capability is an ability to execute routines and solve problems. An organization's capabilities consist of its stock of solutions to the problem of producing collective action in a specified environment. In other words, capabilities are context specific. Capabilities are often based on routines that codify an organization's dispersed learning (March 1988; Nelson and Winter 1982). An important dimension of capability involves the capacity to reduce friction among the many activities and

routines that typically must be undertaken to produce the organization's collective product. The more refined and harmonized an organization's routines, the greater the organization's capability in the specified environment. Therefore, in a stable environment, improvements in capability increase the expected quality of performance and thereby decrease the risk of mortality.

A key part of Stinchcombe's argument for the liability of newness concerns the effect of aging on capabilities. In particular, he argues that new organizations suffer from low average quality of performance because they lack experience. As youthful organizations age, they acquire experience and can potentially learn. The stylized image of the organizational learning curve captures this kind of process. As organizations learn from experience, they refine their productive routines and the metaroutines that coordinate them.

Stinchcombe also argues that new organizations face jeopardy because they must rely on the cooperation of strangers. To the extent that trust enhances collective action, lack of familiarity among coworkers is problematic. As time passes, trust tends to develop within work groups. As a result, the organization's capabilities improve, because an important source of friction has been reduced.

Subsequent theory has followed Stinchcombe's lead in emphasizing that experience improves capabilities. For instance, Hannan and Freeman (1984) argue that norms of rationality demand that organizations achieve low variance in the quality of their outputs and make systematic, rational accounts of their activities. Not all organizations can achieve reliability and accountability; those that do have survival advantages. New organizations lack these capabilities, which must be acquired through learning by doing and the accumulation of organization-specific human capital. Therefore, the development and refinement of these capabilities depends upon age.

Some recent lines of argument about capabilities run opposite this mainstream view. Barron et al. (1994) draw on an analogy to senescence processes observable in animal and human life histories. They suggest that organizations accumulate durable features, such as precedents, political coalitions, and taken-for-granted understandings, that constrain modifications in patterns of collective action. Such encrustation erodes the capability for efficient collective action. According to this view, the liability is one of senescence: mortality rates increase with age.

March's (1991) account of organizational learning tells a cautionary tale about the consequences of the continual refinement of a competence. Organizations that seek to exploit their competencies by searching for ever-better refinements of their existing capabilities find themselves in a compe-

tency trap if the world changes. Only organizations that have already achieved some competence and follow a so-called exploitation strategy can be trapped by their competence. Thus this line of argument opposes the mainstream account, at least in the case of the exploitation strategy.⁵

D. Positional Advantage

A final set of relevant arguments about the effect of age on mortality rates concerns the organization's position in the social structure. Stinchcombe (1965) points out that trust matters in building ties with other organizations (e.g., potential suppliers and customers) and important actors in the social environment (e.g., holders of capital, government officials and regulators, and so forth) and that it takes time for trust to develop in external relations. Because maintaining good relations with key external actors enhances organizational performance and also arguably affects survival chances directly, the buildup of favorable external ties over organizational lifetimes lowers mortality rates. This argument too implies *negative* age dependence in mortality rates.

Positional advantages come in many forms. These include occupancy of positions that bridge structural holes (Burt 1992), favorable reputation (Kreps 1996), high status (Podolny 1993; Podolny, Stuart, and Hannan 1996), market power, and political influence. Some positional advantages such as bridging structural holes would seem not to depend upon age and experience. Others, such as market power and, possibly, political influence, appear to be more influenced by size than by age. Still others, such as favorable reputation and high status, presumably depend upon some demonstrated history of performance and thus depend in part upon age. New organizations thus find themselves at a disadvantage when reputation and status matter. Although not all kinds of positional advantages accrue to experience, I am not aware of any argument claiming that such advantages decline systematically with age. Thus, arguments about position advantage, if they bear on the issue at all, concur with Stinchcombe's original assertion of negative age dependence.

These arguments about endowments, imprinting, positional advantages, and capabilities have been invoked in varying combinations to ex-

⁵ The polar opposite strategy, exploring for solutions that lie far from the organization's current capabilities, has no obvious relationship with aging, as long as the region to be searched is large relative to an organization's radius of search (so that an old organization would not have searched most of the solution space) or the set of possibilities to be searched changes over time.

plain patterns of positive, negative, and nonmonotonic age dependence. Endowment theories are seen to imply positive age dependence. Theories about imprinting effects also appear to imply positive age dependence. Arguments about capabilities can imply either negative or positive age dependence, depending on whether the relevant capabilities are thought to improve or to deteriorate with age. And, it is usually claimed that the cumulation of positional advantages as organizations age results in negative age dependence. Finally, analysts frequently use a mixture of two or more processes to make predictions. For instance, those who emphasize the role of endowments in shaping early mortality experiences have overlaid this process of positive age dependence with standard stories about the age-dependent accumulation of competence and positional advantage. Thus Brüderl and Schüssler (1990), Fichman and Levinthal (1991), and Brüderl et al. (1996) assume that mortality rates rise with age initially due to the exhaustion of endowments and then fall with further aging due to the improvement of capabilities with age. Of course, other combinations are possible.

With the key concepts in hand, we turn now to the formalization. This formalization addresses some of the central ideas about the relationship between age and the rate of organizational mortality. It concentrates on endowment, capability, and position. The formalization does not consider either the effect of size on mortality rates or the interaction of size and age in affecting mortality rates. Doing so would require a substantially more extensive formalization.

III. FORMALIZING THE LIABILITIES OF NEWNESS, ADOLESCENCE, AND SENESENCE

As noted at the outset, this analysis uses the tools of FOL to rationally reconstruct the various theories and to verify the soundness of the derivations. Proofs of all lemmas (minor theorems) and the main theorems were verified with OTTER (Organized Techniques for Theorem-proving and Effective Research; McCune 1994), which is a theorem prover for FOL with the equality relation.

FOL contains the quantifiers: \forall , the universal quantifier ("for all") and \exists , the existential quantifier ("there exists") and the connectives: \neg (not), \wedge (and), \vee (or), \rightarrow (implies) and \leftrightarrow (if and only if). The equality symbol $=$ is a special two-place predicate with the fixed interpretation that a sentence of the form $x = y$ is true if and only if x and y denote the same object.

The notation is summarized in table 1. Uppercase strings denote predicates and functions. Lowercase letters denote the objects that possess the

TABLE 1

NOTATION

Expression	Meaning
Predicate/function:	
$O(x)$	x is an organization
$A(x, t)$	Organization x 's age at time t
$H(x, t)$	Organization x 's hazard of mortality at t
$EN(x)$	Organization x has an endowment at founding
$C(x, t)$	Organization x 's capability at time t
$P(x, t)$	The quality of organization x 's position at t
$IM(x, t)$	Organization x has immunity at t
$K(x, t)$	Organization x 's stock of knowledge at t
$T(x, t)$	The quality of organization x 's external ties at t
$F(x, t)$	Organization x 's level of internal friction at t
$DS(x_0, t)$	The environments at t_0 and t are dissimilar for organization x
$AL(x, t)$	Organization x is aligned with its environment at t
$PA(x, t)$	Organization x has positional advantage at t
$FG(x)$	Organization x occupies a fragile position
$RB(x)$	Organization x occupies a robust position
Parameter:	
ϵ	The age at which an endowment ends
σ	The age at which environmental drift destroys alignment
τ	The age at which advantage begins to accrue to robust position

properties or are mapped by the functions. One-place predicates denote the properties of objects. For instance, the predicate $O(x)$ indicates that an object x is an organization. Predicates with more than one argument "slot" and variables denote relations between objects. (Predicates tell that a certain relationship is true; functions define mappings of objects.) For instance, the basic variable in these formalizations is historical time, denoted by t . The function of most substantive relevance, $A(x, t)$, gives an object's age at time t .

A. Capability, Position, and Mortality

This subsection presents a set of propositions that—in one form or another—find application in theories of the liability of newness, adolescence, senescence, and obsolescence. These substantive arguments concern the effects of organizational age on the hazard of mortality. Two properties of the age function deserve note: age is nonnegative and increases monotonically with time in existence:

MEANING POSTULATE 1.—*Age is nonnegative and monotonic in t .*

$$\forall x, t_0, t [O(x) \rightarrow [A(x, t) \geq 0]$$

$$\wedge [[A(x, t) > A(x, t_0) \rightarrow (t > t_0)] \wedge [A(x, t) = A(x, t_0) \rightarrow t = t_0]].$$

Read: for all values of x , t_0 and t , if an object x is an (existing) organization, then the object's age is greater than or equal to zero, and the statement that the object's age at time t exceeds its age at t_0 implies that time t is greater than time t_0 , and the statement that the object's age at t equals its age at t_0 implies that $t = t_0$.

The dependent variable, $H(x, t)$, is an organization's instantaneous rate of mortality at time t . In some earlier formalizations (Péli et al. 1994; Péli 1997), the mortality rate has been interpreted as a population-level outflow. In the present context, the rate should be interpreted as an organization-level hazard. (The appendix contains a pair of technical background assumptions used in deriving results about effects on the hazard.)

The substantive core for the first three formalizations involves the relations of endowments, capability, position, and the hazard of mortality. The review of the arguments about endowment effects in organizational mortality suggested that endowments play the role of transitory initial advantages. They buffer new organizations from some of the force of selection processes for a time. But these advantages get exhausted relatively quickly (in the scale of organizational lifetimes). On this interpretation, endowments can be conceptualized as transitory immunity from the risk of mortality. Suppose that an endowment lasts until age ϵ . This idea can be represented in terms of a predicate that tells that an organization possesses immunity at a given time: $IM(x, t)$. That is, an endowment provides immunity that lasts from founding until age ϵ (meaning postulate 1, which states that age is nonnegative, requires that $\epsilon \geq 0$; however, the case when $\epsilon = 0$ is uninteresting, because there is no effect of the endowment).

DEFINITION 1.—*An endowment provides an immunity that lasts until an organization's age exceeds ϵ .*

$$\begin{aligned} \forall x(EN(x) \leftrightarrow \forall t [O(x) \wedge [A(x, t) \leq \epsilon] \rightarrow IM(x, t)] \\ \wedge [A(x, t) > \epsilon] \rightarrow \neg IM(x, t)]). \end{aligned}$$

In the context, it makes sense to specify that endowments are the only source of immunity.

ASSUMPTION 1.—*An unendowed organization never possesses immunity.*

$$\forall x, t [O(x) \wedge \neg EN(x) \rightarrow \neg IM(x, t)].$$

Finally, we need a specification of the role of immunity in the process of organizational mortality. According to the image used here, the hazard of mortality remains constant at a very low level (possibly zero) during the period of immunity.⁶ There are two ideas here. First, the hazard stays constant during the period of immunity.

ASSUMPTION 2.—*An organization's hazard of mortality is constant during periods in which it has immunity.*

$$\forall x, t_0, t [O(x) \wedge IM(x, t_0) \wedge IM(x, t) \rightarrow H(x, t_0) = H(x, t)].$$

Second, immunity has an overwhelming effect on the mortality rate, an effect that overrides the effects of other potentially relevant conditions such as capability and quality of position.

ASSUMPTION 3.—*An organization's hazard of mortality is lower during periods in which it has immunity than in periods in which it does not.*

$$\forall x, t_0, t [O(x) \wedge IM(x, t_0) \wedge \neg IM(x, t) \rightarrow H(x, t) > H(x, t_0)].$$

Outside of periods of immunity, variations in capability and performance affect the hazard. The function $C(x, t)$ maps an object's capability at time t and $P(x, t)$ tells the quality of its position at time t .

ASSUMPTION 4.—*When an organization lacks immunity, superior capability and position imply a lower hazard of mortality.*

$$\begin{aligned} \forall x, t_0, t [O(x) \wedge \neg [IM(x, t_0) \vee IM(x, t)] \\ \rightarrow \{ [C(x, t) > C(x, t_0) \wedge P(x, t) \geq P(x, t_0)] \\ \vee [C(x, t) \geq C(x, t_0) \wedge P(x, t) > P(x, t_0)] \rightarrow H(x, t) < H(x, t_0) \} \\ \wedge \{ [C(x, t) = C(x, t_0) \wedge P(x, t) = P(x, t_0)] \rightarrow H(x, t) = H(x, t_0) \}). \end{aligned}$$

The liability-of-newness theory builds upon assumptions that relate an organization's age to the development of its capabilities and positional advantages. Stinchcombe and others have identified several processes by which aging enhances capability and position in a constant environment. To simplify, this formalization considers the stock of organizational knowledge as an instance of capability and the quality of ties to external actors who control important resources as an instance of the quality of a structural position.⁷ The function $K(x, t)$ records an object's collective

⁶ Another, more complicated possibility is that an immunity wears off gradually, so that the hazard of mortality rises gradually during the immunity period. Such a complication does not appear to be needed here.

⁷ It should be obvious that a complete formulation would develop the forms of capabilities and positions more fully.

knowledge at time t ; and $T(x, t)$ tells the quality of its external ties at time t .

The main intuition behind the senescence argument holds that organizations develop encrustations of internal frictions, precedents, and political compromises as they age and that these encrustations impede timely collective action. A diminished capacity for timely collective action amounts to a loss of capability. This formalization also considers only one instance of the purportedly relevant conditions: the accumulation of internal friction. The function $F(x, t)$ records an object's intensity of internal friction at time t .

According to the standard liability-of-newness story, collective knowledge contributes to capability. Arguments for senescence assume that the accumulation of internal friction diminishes capability. It is helpful to develop a simple framework that applies to both cases so that each theory can be regarded as a special case of the general formulation. The liability-of-newness proposition will be derived by assuming that internal friction does not vary with age while capability and position increase with age; and the liability-of-senescence proposition will be derived by assuming that internal friction accumulates with age while capability and position remain constant. The core assumptions presented here are tailored to fit these requirements.⁸

ASSUMPTION 5.—*Increased knowledge elevates an organization's capability; and increased accumulation of organizational internal frictions diminishes its capability.*

$\forall x, t_0, t (O(x))$

$$\begin{aligned} &\rightarrow [[K(x, t) > K(x, t_0)] \wedge [F(x, t) \leq F(x, t_0)] \rightarrow C(x, t) > C(x, t_0)] \\ &\wedge [[K(x, t) \leq K(x, t_0)] \wedge [F(x, t) > F(x, t_0)] \rightarrow C(x, t) < C(x, t_0)] \\ &[[K(x, t) = K(x, t_0)] \wedge [F(x, t) = F(x, t_0)] \rightarrow C(x, t) = C(x, t_0)]. \end{aligned}$$

ASSUMPTION 6.—*Improved ties with external actors enhance an organization's position.*

$$\begin{aligned} &\forall x, t_0, t (O(x) \rightarrow \{[T(x, t) > T(x, t_0)] \rightarrow P(x, t) > P(x, t_0)\} \\ &\wedge [T(x, t) = T(x, t_0)] \rightarrow P(x, t) = P(x, t_0)\}). \end{aligned}$$

With these assumptions in place, mortality hazards can be related to knowledge, internal friction, and ties (see fig. 1). The following pair of lemmas establishes the connections. Note that they apply only to periods

⁸ It would be valuable to develop a more complete formalization of these issues; but doing so here would detract from the main line of argument and formalization.

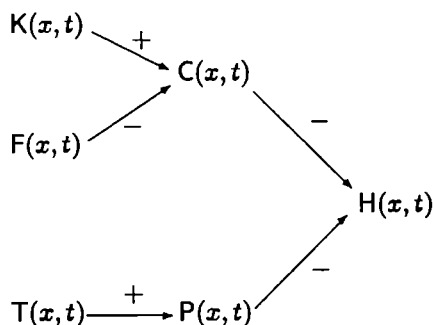


FIG. 1.—A schematic representation of the expanded version of the core for organizations lacking immunity. The terms are knowledge, $K(x, t)$; ties, $T(x, t)$; friction, $F(x, t)$; capability, $C(x, t)$; position, $P(x, t)$; and the hazard of mortality, $H(x, t)$.

in which organizations lack immunity (either because an organization lacks endowment or because the period of immunity produced by an endowment has ended.) During periods of immunity, the hazard is a constant according to assumption 2.

LEMMA 1.—*When an organization lacks immunity, increased collective knowledge and superior external ties lower its hazard of mortality when internal friction does not increase (from assumptions 1–6).*

$$\begin{aligned}
 &\forall x, t_0, t \{O(x) \wedge \neg [IM(x, t_0) \vee IM(x, t)] \\
 &\quad \rightarrow [(K(x, t) > K(x, t_0) \wedge F(x, t) \leq F(x, t_0) \wedge T(x, t) \geq T(x, t_0)) \\
 &\quad \vee [K(x, t) \geq K(x, t_0) \wedge F(x, t) \leq F(x, t_0) \wedge T(x, t) > T(x, t_0)] \\
 &\quad \rightarrow H(x, t) < H(x, t_0)]\}.
 \end{aligned}$$

LEMMA 2.—*When an organization lacks immunity, the growth of internal friction elevates its hazard of mortality when its knowledge and the quality of its ties are constant (from assumptions 1–6).*

$$\begin{aligned}
 &\forall x, t_0, t \{O(x) \wedge \neg [IM(x, t_0) \vee IM(x, t)] \wedge K(x, t) = K(x, t_0) \\
 &\quad \wedge F(x, t) > F(x, t_0) \wedge T(x, t_0) = T(x, t) \rightarrow H(x, t) > H(x, t_0)\}.
 \end{aligned}$$

B. The Liabilities of Newness and Adolescence

The liability-of-newness argument, as formalized here, rests on the premise that the processes that improve capability and position—knowledge and the quality of ties—increase monotonically with organizational age:

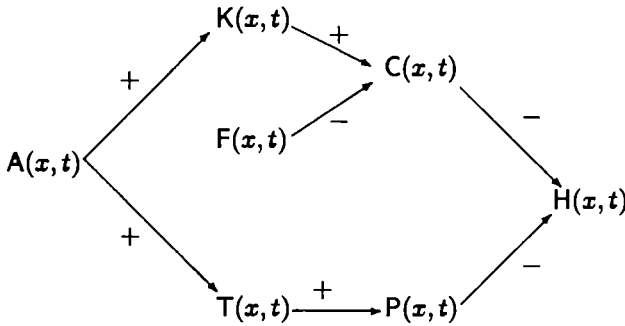


FIG. 2.—A schematic representation of effect of age, $A(x, t)$, for organizations lacking immunity, according to the liability of newness argument.

ASSUMPTION 7.—*An organization's stock of knowledge increases monotonically with its age.*

$$\forall x, t_0, t [O(x) \wedge A(x, t) > A(x, t_0) \rightarrow K(x, t) > K(x, t_0)].$$

ASSUMPTION 8.—*The quality of an organization's external ties increases monotonically with its age.*

$$\forall x, t_0, t [O(x) \wedge A(x, t) > A(x, t_0) \rightarrow T(x, t) > T(x, t_0)].$$

The standard argument for the liability of newness does not consider the effect of internal friction. For the formalization to remain faithful to the original argument, it should not mix this process with the others. To meet this constraint, this section assumes that an organization's internal friction remains constant over age.

ASSUMPTION 9.—*An organization's internal friction does not vary with its age.*

$$\forall x, t_0, t [O(x) \rightarrow F(x, t) = F(x, t_0)].$$

These assumptions yield straightforward links between age and capability and quality of position (see fig. 2). The following lemmas record the implications of these assumptions for age variations in capability and position.

LEMMA 3.—*An organization's capability increases monotonically with its age (from assumptions 5 and 7–9).*

$$\forall x, t_0, t [O(x) \wedge A(x, t) > A(x, t_0) \rightarrow C(x, t) > C(x, t_0)].$$

LEMMA 4.—*An organization's structural position improves monotonically with its age (from assumptions 8 and 6).*

$$\forall x, t_0, t [O(x) \wedge A(x, t) > A(x, t_0) \rightarrow P(x, t) > P(x, t_0)].$$

The foregoing definitions, assumptions, and lemmas imply Stinchcombe's theorem on the liability of newness as a consequence of age-dependent enhancement of capability and position for the case of organizations lacking endowments:

THEOREM 1.—*The liability-of-newness theorem (Stinchcombe 1965): an unendowed organization's hazard of mortality declines monotonically with its age (from definition 1, assumptions 1–4, and lemmas 1 and 2).*

$$\forall x, t_0, t [O(x) \wedge \neg EN(x) \wedge A(x, t) > A(x, t_0) \rightarrow H(x, t) < H(x, t_0)].$$

This formulation also implies a theorem on the liability of adolescence as a consequence of age-dependent enhancement of capability and position for the case of organizations with endowments.

THEOREM 2.—*The liability-of-adolescence theorem (Brüderl and Schüssler 1990; Fichman and Levinthal 1991): an endowed organization's hazard of mortality is constant during its period of immunity, jumps when its immunity ends, and decreases with further aging but remains above the level during the immunity period (from definition 1, assumptions 1–4, and lemmas 1 and 2).*

$$\begin{aligned} \forall x, t_0, t_1, t_2, t_3 [O(x) \wedge EN(x) \wedge [A(x, t_0) \leq A(x, t_1)] \\ \wedge [A(x, t_1) \leq \epsilon] \wedge [A(x, t_2) > \epsilon] \wedge [A(x, t_3) > A(x, t_2)] \\ \rightarrow H(x, t_2) > H(x, t_3) > H(x, t_1) = H(x, t_0)]. \end{aligned}$$

C. The Liability of Senescence

Next consider the first of the opposing arguments: the senescence hypothesis. The formalization in this section uses the same general structure as discussed to this point. However, it makes crucial changes in the three assumptions pertaining to the effect of aging on the determinants of organizational performance, assumptions 7–9. These three assumptions are replaced with others that reflect the senescence argument as stated in Section II.

The senescence argument emphasizes capability. It identifies processes that diminish capability and argues that these processes increase in strength as organizations age. Here this effect is represented by the accumulation of internal frictions. The spirit of the senescence argument suggests that one assume that internal friction increases with age.

Extant statements of the senescence story are silent about the role of knowledge and structural position. Therefore, to represent this argument, assume now that knowledge and the quality of ties to external actors also do not change with age.

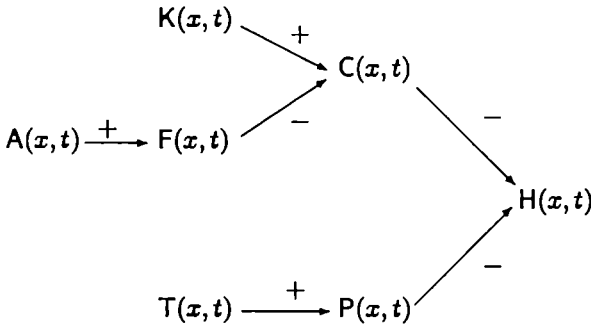


FIG. 3.—A schematic representation of the liability of senescence argument for organizations lacking immunity.

ASSUMPTION 10.—*An organization's stock of organizational knowledge does not vary with its age* (contra assumption 7).

$$\forall x, t_0, t [O(x) \rightarrow K(x, t) = K(x, t_0)].$$

ASSUMPTION 11.—*The quality of an organization's external ties does not vary with its age* (contra assumption 8).

$$\forall x, t_0, t [O(x) \rightarrow T(x, t) = T(x, t_0)].$$

ASSUMPTION 12.—*An organization's internal friction increases monotonically with its age* (revising assumption 9).

$$\forall x, t_0, t [O(x) \wedge A(x, t) > A(x, t_0) \rightarrow F(x, t) > F(x, t_0)].$$

With these assumptions about knowledge, position, and friction, only the strength of internal friction depends upon age (see fig. 3). These assumptions portray a situation in which an organizations grows older without growing wiser and accumulates baggage that impedes its collective action. The following pair of (trivial) lemmas connects these assumptions to the hazard.

LEMMA 5.—*An organization's capability decreases monotonically with its age* (from assumptions 5, 10, and 12).

$$\forall x, t_0, t [O(x) \wedge A(x, t) > A(x, t_0) \rightarrow C(x, t) > C(x, t_0)].$$

LEMMA 6.—*An organization's structural position does not vary with its age* (from assumptions 6 and 9).

$$\forall x, t_0, t [O(x) \wedge A(x, t) > A(x, t_0) \rightarrow P(x, t) = P(x, t_0)].$$

The assumptions in this section imply *positive* age dependence in mortality hazard for an unendowed organization (contra theorem 1).

THEOREM 3.—*The senescence theorem for unendowed organizations (Barron et al. 1994): an unendowed organization's hazard of mortality increases monotonically with its age* (from definition 1, assumptions 1–4, and lemmas 5 and 6).

$$\forall x, t_0, t [O(x) \wedge \neg EN(x) \wedge A(x, t) > A(x, t_0) \rightarrow H(x, t) > H(x, t_0)].$$

Previous work did not consider the potential effect of endowments in the context of senescence. The formulation in this section carries over to the case of endowed organizations. There is, however, a slight difference: due to initial immunity, the hazard does not rise monotonically at all ages for endowed organizations. Instead, the hazard is constant during the period of immunity and then rises monotonically with age.

THEOREM 4.—*The senescence theorem for endowed organizations: An endowed organization's hazard of mortality remains constant during the period of immunity and increases monotonically with its age once immunity ends* (from definition 1, assumptions 1–4, and lemmas 5 and 6).

$$\begin{aligned} \forall x, t_0, t_1, t_2, t_3 [O(x) \wedge EN(x) \wedge [A(x, t_0) \leq A(x, t_1)] \\ \wedge [A(x, t_1) \leq \epsilon] \wedge [A(x, t_2) > \epsilon] \wedge [A(x, t_2) > A(x, t_1)] \\ \rightarrow H(x, t_3) > H(x, t_2) > H(x, t_1) = H(x, t_0)]. \end{aligned}$$

The reasoning underlying the claims of liabilities of newness and senescence run parallel, with differing assumptions about the role of age in shaping capability and position. Each story can be regarded as a special case of a more general story in which aging can both enhance and diminish capability, depending on the relative strength of the links of age with the cumulation of knowledge and the intensification of internal friction.

IV. FORMALIZING ALIGNMENT, ENVIRONMENTAL DRIFT, AND THE LIABILITY OF OBSOLESCENCE

The remainder of the formalization addresses the implications of imprinting and obsolescence. Making this shift involves some changes in the formalisms. The two theories of age dependence considered in the previous section fit nicely with statements made in terms of functions because the theoretical notions concern processes that increase continuously with age. The obsolescence argument builds on the image of some things (environments) shifting and other things (core features of organizations) getting stuck, and it seems best to represent this argument with predicates that tell whether certain objects have shifted and whether others have not. The remainder of the article presents two kinds of formalizations of the

obsolescence process. The first version closely parallels the formalizations just discussed, using functions such as capability and age. The second formalization relies extensively on predicates. Because the modeling structure in both cases differs from that used in the previous section, this section begins (nearly) anew. Several features of the core of the foregoing formalization carry over. These are meaning postulates 1 (age is nonnegative) and 2 (age is monotonic) and the definitions of the functions and predicates. The definition of endowment as providing temporary immunity (definition 1) and three assumptions about endowments (assumption 1, unendowed organizations lack immunity; assumption 2, the hazard is constant within periods of immunity; and assumption 3, the hazard is lower during periods of immunity) are also retained.

A. An Alternative Core: Alignment in Drifting Environments

The theory to be addressed in this section concerns the effects of the combination of environmental change and inertia rather than genuine aging. The key modeling issue here is how to formalize the relevant part of the imprinting hypothesis. One approach, which is used here, considers external alignment, the fit between an organization's structural features and capabilities and the demands of its external environment (Hannan 1996). The obsolescence story posits that *capabilities are specific to environments*.⁹ Refining obsolete routines (by, e.g., learning by doing) and better coordinating activities that support obsolete routines do not have mortality-diminishing effects. In other words, external misalignment can be regarded as a mismatch of capabilities and environments. This idea is represented by the predicate, $AL(x, t)$, which tells whether an object, x , is (or is not) aligned with its environment at time t . That is, an organization finds itself either aligned or not aligned with its environment.

Next, consider the issue of environmental change, which drives the obsolescence process. Suppose that the environment can occupy different states at different times, meaning that it imposes different adaptive demands at different times. Does it make sense to think of an organization as being aligned with more than one state of the environment? If the environments impose similar adaptive demands, it does. But the notion of alignment also suggests that an organization's capability cannot be aligned with dissimilar states of the environment. In other words, dissimilarity can be defined in terms of alignment. Two states of the environment

⁹ This idea seems implicit in the theoretical treatments of the liability of newness. However, its implications seem not to have been appreciated.

impose dissimilar adaptive demands if an organization cannot be aligned with both.

This idea can be represented with the three-place predicate, $DS(x, t_0, t)$, which tells whether the environments at times t_0 and t are dissimilar for organization x .¹⁰

DEFINITION 2.—*Two states of the environment are dissimilar if and only if organizations cannot be aligned with each.*

$$\forall t_0, t (DS(x, t_0, t) \leftrightarrow \{O(x) \wedge \neg [AL(x, t_0) \leftrightarrow AL(x, t)]\}).$$

In substantive context, this assumption means that the alternative states of the environment pose very different adaptive demands. Therefore, this formulation applies to situations in which the states of the environment differ greatly when compared with the repertoires of organizational routines available to the members of the population of organizations. Two properties of the dissimilarity predicate need to be noted. First, there is a piece of "background knowledge" that tells that any environment is similar with itself.

BACKGROUND ASSUMPTION 1.—*The environment at any time is similar with itself.*

$$\forall t [\neg DS(x, t, t)].$$

Second, background assumption 1 and definition 2 imply

LEMMA 7.—*The dissimilarity relation is symmetric.*

$$\forall t_1, t_2 [DS(x, t_1, t_2) \leftrightarrow DS(t_2, t_1)].$$

A strong-form version of the imprinting hypothesis posits that alignment is maximal at (or shortly after) founding, because organization-builders can make use of state-of-the art designs and adapt to prevailing cultural understandings (Hannan, Burton, and Baron 1996). That is, designers can tune structures and routines to the state of the environment at the time of founding. If the environment changes, then the inertial forces on structures, processes, and capabilities restrict the ability of organizations to re-align with the new environment.¹¹

¹⁰ Obviously some rule must be used in practice to determine how different two environments must be for them to be considered "dissimilar." One could replace the predicate with a function that tells the degree of dissimilarity of the two environments. However, this choice would make the formalization more complex; it is not needed to convey the main idea.

¹¹ Of course, many organizations begin with poor designs and poor alignment with even the founding environment. Unobserved heterogeneity in initial design and initial alignment is a well-understood source of spurious negative age dependence (of the mover-stayer form, see Freeman et al. 1983). To keep the focus on the role of age, this article does not consider such misalignment.

ASSUMPTION 13.—*An organisation is aligned with the state of the environment at its time of founding.*

$$\forall x, t \{O(x) \wedge [A(x, t) = 0] \rightarrow AL(x, t)\}.$$

According to theories of alignment, organizations have superior capabilities in environments with which they align.

ASSUMPTION 14.—*An organisation's capability is higher in the state of the environment to which it is aligned.*

$$\forall x, t_0, t [O(x) \wedge AL(x, t_0) \wedge DS(x, t_0, t) \rightarrow C(x, t_0) > C(x, t)].$$

Representing the idea of obsolescence requires specification of a time path of change in the relevant environment. According to the general idea of obsolescence, organizations get stuck with core features that become ill-suited to the environment with the passage of time. Although many different scenarios of environmental change are consistent with this image, this article considers only one simple case that seems to capture the key insight: *drift*. If an environment drifts, its states at (widely) different times are (strongly) dissimilar. When an environment drifts, the probability that it returns to some previous state is vanishingly small. (In the formalization, this event is assumed to be impossible.) For instance, technological change involves drift: technical standards change over time and do not return to earlier standards.

Environmental drift differs from the broader category of environmental change. Environmental change can involve repeated shifts among alternative states with the possibility of returns to earlier states. For instance, sociological models of niche width (Freeman and Hannan 1983; Péli 1997) consider cases in which environments can alternate between two states, for example, feast or famine. When an organization or population of organizations finds itself misaligned with the current state of such an environment, it should not be regarded as "obsolete." The concept of obsolescence seems most meaningful in the case of drifting environments. The remainder of the formalization specializes to the case of drifting environments.

The notion of drift means that an organization's environment might remain in the close neighborhood of a given position for some time but will eventually move beyond the neighborhood. Suppose the neighborhood boundary encloses the set of similar environmental states (from the perspective of the organization). Then, as long as the environment stays within the neighborhood, the various states of the environment are similar. Suppose that a drifting environment does not return to its initial neighborhood once it leaves and that there is a (common) age, σ , at which the environment drifts beyond a given initial neighborhood. Then, age and dissimilarity are linked:

ASSUMPTION 15.—*Environmental drift: the environments at times separated by more than σ are dissimilar.*

$$\forall x, t_0, t \ (O(x) \wedge [A(x, t_0) = 0] \rightarrow [[A(x, t) > \sigma] \leftrightarrow DS(x, t_0, t)]).$$

B. The Liability of Obsolescence in Drifting Environments

Existing statements of the obsolescence argument consider capability but not positional advantage. To represent these arguments here, I continue to assume that quality of position does not play a role, by weakening lemma 1 from the previous formulation to apply only to capability and mortality hazards and implementing the lemma as an assumption in this formulation:

ASSUMPTION 16.—*Superiority in capability lowers the hazard of mortality when an organization lacks immunity.*

$$\begin{aligned} \forall x, t_0, t \ \{ & O(x) \wedge \neg [IM(x, t_0) \vee IM(x, t)] \\ & \wedge C(x, t) > C(x, t_0) \rightarrow H(x, t_0) > H(x, t) \}. \end{aligned}$$

Combining the assumptions of environmental drift and constancy of position with the assumptions in Section IVA yields two theorems about obsolescence. One concerns unendowed organizations, and the second concerns endowed organizations. Both tell that mortality rates of old organizations exceed those of young ones. The only difference is the profile of positive age dependence. Age dependence is monotonic positive over the whole age range for unendowed organizations. For endowed organizations, the hazard remains constant during the period of immunity and then rises monotonically with age when immunity ends.

THEOREM 5.—*The obsolescence theorem for unendowed organizations (Barron et al. 1994): an unendowed organization's hazard of mortality increases with age in a drifting environment* (from definitions 1 and 2, background assumption 1, and assumptions 1–3, 13–15, and 16).

$$\begin{aligned} \forall x, t_0, t_1, t_2 \ \{ & O(x) \wedge \neg EN(x) \wedge [A(x, t_0) = 0] \wedge [A(x, t_1) \leq \sigma] \\ & \wedge [A(x, t_2) > \sigma] \wedge (\sigma > 0) \rightarrow H(x, t_2) > H(x, t_1) \}. \end{aligned}$$

Again a parallel theorem holds for endowed organizations. However, this case involves a slight complication because two parameters, ϵ and σ , matter. The following pair of lemmas cover the two relevant cases ($\epsilon \geq \sigma$ and $\epsilon < \sigma$).

LEMMA 8.—*When $\epsilon < \sigma$ in a drifting environment, an endowed organization's hazard of mortality remains constant until age reaches ϵ , then jumps to a higher level, then jumps again at age σ* (from definitions 1 and 2, background assumption 1, and assumptions 1–3, 13–15, and 16).

$$\begin{aligned} \forall x, t_0, t_1, t_2, t_3 \{ & O(x) \wedge EN(x) \wedge [A(x, t_0) = 0] \wedge [A(x, t_1) \leq \epsilon] \wedge [A(x, t_2) > \epsilon] \\ & \wedge [A(x, t_2) \leq \sigma] \wedge [A(x, t_3) > \sigma] \wedge (\sigma > \epsilon) \wedge (\epsilon > 0) \\ & \rightarrow H(x, t_3) > H(x, t_2) > H(x, t_1) = H(x, t_0) \}. \end{aligned}$$

LEMMA 9.—*When $\epsilon \geq \sigma$, an endowed organization's hazard of mortality remains constant until age ϵ and then jumps to a higher level in a drifting environment* (from definitions 1 and 2, background assumption 1, and assumptions 1–3, 13–15, and 16).

$$\begin{aligned} \forall x, t_0, t_1, t_2 \{ & O(x) \wedge EN(x) \wedge [A(x, t_0) = 0] \wedge [A(x, t_1) \leq \epsilon] \wedge [A(x, t_2) > \epsilon] \\ & \wedge (\epsilon \geq \sigma) \wedge (\sigma > 0) \rightarrow H(x, t_2) > H(x, t_1) = H(x, t_0) \}. \end{aligned}$$

Obviously, the hazard at older ages exceeds the hazard during the youthful (immunized) period.

THEOREM 6.—*The obsolescence theorem for endowed organizations: In a drifting environment, an endowed organization's hazard of mortality is constant during the period of immunity; beyond the period of immunity, the hazard rises with age* (from lemmas 8 and 9).

$$\begin{aligned} \forall x, t_0, t_1, t_2, t_3 \{ & O(x) \wedge EN(x) \wedge [A(x, t_0) = 0] \wedge [A(x, t_1) \leq \epsilon] \\ & \wedge [A(x, t_2) > \epsilon] \wedge (\epsilon > 0) \rightarrow H(x, t_2) > H(x, t_1) = H(x, t_0) \}. \end{aligned}$$

C. Fragile and Robust Positions in Drifting Environments

The formulations considered to this point show that a reasonably general set of assumptions about capability and position can be altered to accommodate four theories of age dependence: those implying liabilities of newness, adolescence, senescence, and obsolescence. These formulations differ about the sign of age dependence—it is negative in the first, has an inverted-U shape in the second, and is positive in the other two. All of these formulations treat a population of organizations as *homogeneous* with respect to the process. Yet, as noted at the outset, some recent research reveals more complex patterns in which the form of age dependence differs for different classes of organizations in a population. In the relevant research (Hannan et al. 1998), the factor separating the classes is organizational size: mortality hazards rise with age for small organizations and decline with age for large ones. Given these results, it is interesting to explore how the assumptions about capability, position, and aging might be altered so as to produce such a pattern. This section takes up this task by reformulating the assumptions about position. It builds on the set of definitions and assumptions in Section IVA but does not use those in Section IVB.

Consider the issue of positional advantage in the case of drifting environments. It seems unreasonable to assume that all positional advantages

can be preserved when environments drift. This section introduces a distinction between two types of structural position: fragile and robust. Fragile positions depend upon some potentially unstable configuration of ties. In this context, instability means that the stability of a position or the advantages accruing to incumbency in a position are very sensitive to the state of the environment—small changes in the environment can destroy the position or its value. For instance, an organization whose positional advantage stems from its ties with a particular political leader occupies a fragile position. If political conditions change and this leader loses power, then the positional advantage vanishes. In contrast, robust positions depend upon the occupation of highly reproducible positions, those whose value is relatively insensitive to the state of the environment. High status provides a good example of a robust position, because a positive feedback loop runs from high status to capability; for example, high-status actors can produce signals of high quality at lower cost (Podolny 1993). Because environmental changes likely damage fragile positions more than robust ones, fragile and robust positions can be defined in terms of the degree to which their advantages generalize over environments.

This formalization builds on the assumption that gaining advantage from robust position takes time but that fragile positions can provide advantages at founding. This assumption appears to fit the examples of fragile and robust positions used above. Consider examples of each case. If a fragile position builds on a favorable tie to a particular actor, there is no reason why this tie cannot provide advantage from an organization's founding (when, e.g., it depends on the ties between the leaders of the organization and the external actor). If, in contrast, organizations building robust positions invest in high quality—which will eventually generate positional advantages such as favorable reputation and high status that have value in many states of the environment—then such a positional investment is unlikely to pay off immediately. For instance, it takes time for quality to be acknowledged and status to be conferred. For simplicity, assume that an organization does not obtain advantage from robust position until it passes age τ .

The formulations considered to this point assume a range of positions with varying quality, represented by the function $P(x, t)$. It helps clarify issues here to narrow and consider only two values, using the predicate $PA(x, t)$, which records the presence/absence of positional advantage. Although organizations might be able change their type of position, this formalization assumes that they cannot.¹¹

¹¹ This assumption is undoubtedly too strong. The key idea is that positions endure, that they change much more slowly than the environmental variables influencing life

We need formal definitions of fragile and robust position that fit these images. For fragile position, the key idea holds that environmental change destroys the value of the position. It will simplify the analysis to tie loss of advantage from fragile positions to the specification of drift already developed. Environmental drift eliminates alignment, and it also wrecks fragile positions. Alignment and fragile position might in general differ in their sensitivity to environmental change. One could incorporate such a difference by introducing another drift parameter that applies to fragile position. Not much would seem to be gained by adding such a complication. This formalization employs the simplifying assumption that the same parameter, σ , governs both the loss of alignment and the loss of positional advantage from fragile position.

DEFINITION 3.—*An organization's position is fragile if and only if it does not provide advantage after age σ .*

$$\forall x (FG(x) \leftrightarrow \{[A(x, t) > \sigma] \rightarrow \neg PA(x, t)\}).$$

According to the image of robust position, drift does not matter. All that needs to be specified is the time that it takes for an organization to begin gaining advantage from occupancy of a robust position. This duration is denoted here as τ .

DEFINITION 4.—*An organization's position is robust if and only if it provides positional advantages only after age τ .*

$$\forall x (FG(x) \leftrightarrow \{[A(x, t) \leq \sigma] \rightarrow PA(x, t)\} \wedge \{[A(x, t) > \sigma] \rightarrow \neg PA(x, t)\}).$$

If an organization's position can be both fragile and robust, then the foregoing definitions yield contradictions. For instance, for the portion of the age range in which age is less than σ and less than τ , definition 3 says that the organization possesses positional advantage and definition 4 says that it does not.

LEMMA 10.—*An organization's position cannot be both fragile and robust (from definitions 3 and 4).*

$$\forall x \neg [FG(x) \wedge RB(x)].$$

Finally, a rule must be stated for combining the effects of capability and positional advantage on mortality hazards. (This section bypasses the link through performance here to simplify the argument.) In the style of Péli (1997), one can introduce *ordered levels* of the hazard of mortality, using the name constants: *very_low*, *low*, *mod₁*, *mod₂*, and *high*.

ASSUMPTION 17.—*An organization's immunity, alignment of capability*

chances. I simplify to the assumption that positions remain constant over lifetimes to avoid overly complicating the story at this point

with the current state of the environment and positional advantage jointly affect the hazard of mortality with the following ordinal scaling:

$$\begin{aligned} \forall x, t \quad & (O(x) \rightarrow [IM(x, t) \rightarrow H(x, t) = \text{very_low}] \wedge [\neg IM(x, t) \\ & \rightarrow [AL(x, t) \wedge PA(x, t) \rightarrow H(x, t) = \text{low}] \\ & \wedge [\neg AL(x, t) \wedge PA(x, t) \rightarrow H(x, t) = \text{mod}_1] \\ & \wedge [AL(x, t) \wedge \neg PA(x, t) \rightarrow H(x, t) = \text{mod}_2] \\ & \wedge [\neg AL(x, t) \wedge \neg PA(x, t) \rightarrow H(x, t) = \text{high}]]). \end{aligned}$$

Derivations require that the inequalities implied by the natural language hold in the formalization for the name constants denoting the levels of the mortality hazards:

ASSUMPTION 18.—*The levels of the hazard of mortality are ordered:*

$$\text{high} > \text{mod}_1 > \text{low} > \text{very_low} \wedge \text{high} > \text{mod}_2 > \text{low} > \text{very_low}.$$

Note that assumptions 18 and 19 (below) implement the ideas that immunity overrides other potential forces affecting the hazard and that the hazard remains constant during periods of immunity. Thus these assumptions have the same effect as assumptions 2 and 3 in the formalization used in Section III.

Determinate results also depend on some specification of the relative importance of capability (alignment) and position. This formalization takes the view that positional advantage matters more than the alignment of capability with the environment for the hazard of mortality ($\text{mod}_2 > \text{mod}_1$). This assumption reflects sociological arguments about the effects of position on the returns to quality performance. For instance, Podolny (1993) shows that a "Matthew effect" applies to organizations: organizations that occupy high status positions get more credit than those that occupy low status positions for a performance of a given quality. Likewise Burt's (1992) arguments about "structural holes" also imply that positional advantage dominates. This assumption is implemented by refining this ordinal scaling of the hazards.

ASSUMPTION 19.—*Position dominates alignment:*

$$\text{mod}_2 > \text{mod}_1.$$

It turns out that this assumption does not matter substantively in most of what follows. That is, most of the results also follow if the inequality in this assumption is reversed, because the results usually depend upon a contrast of either of the *mod* categories with *very_low* or *low* or with *high*. Cases in which this assumption plays an important role will be noted in what follows.

What do these assumptions imply about age dependence in mortality

TABLE 2

CASE 1

		ROBUST POSITION	
		$A \leq \tau$	$A > \tau$
$A \leq \sigma$	FRAGILE POSITION	$AL \wedge \neg PA \rightarrow$ $H = \text{low}$	$AL \wedge PA \rightarrow$ $H = \text{low}$
		$AL \wedge \neg PA \rightarrow$ $H = \text{mod}_1$	$AL \wedge PA \rightarrow$ $H = \text{mod}_1$
$A > \sigma$	FRAGILE POSITION	$\neg(AL \vee PA) \rightarrow$ $H = \text{high}$	$\neg(AL \vee PA) \rightarrow$ $H = \text{high}$
		$\neg(AL \vee PA) \rightarrow$ $H = \text{mod}_1$	$\neg(AL \vee PA) \rightarrow$ $H = \text{mod}_1$

NOTE.—These are the implications of drift for the hazard of mortality for organizations lacking immunity with robust and fragile positions

hazards? Answering this question requires that one “calculate” alignment and positional advantage by cuts on the age scale for organizations with fragile and robust positions with and without endowments.

Case 1: organizations lacking endowments.—Consider first the simpler case of organizations lacking endowments. The main result is that the pattern of age dependence in the hazard of mortality diverges by type of position. The top panel of table 2 gives the results, and figure 4 displays the implied pattern of age dependence for unendowed organizations with fragile position. (Note that τ is irrelevant in this case.) The hazard of mortality rises with age:

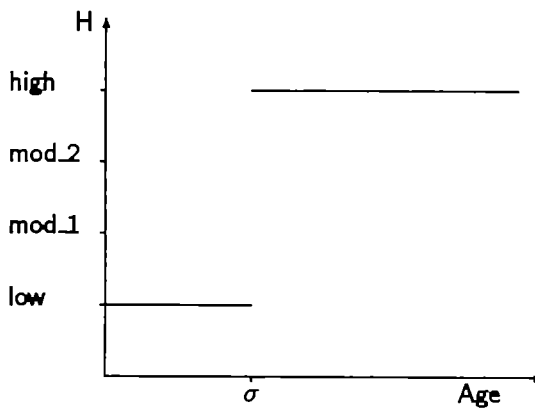


FIG. 4.—The implied pattern of age dependence in mortality hazards for unendowed organizations with fragile positions in drifting environments.

THEOREM 7.—*Fragile position without endowment: The hazard of mortality increases with age for an unendowed organization with a fragile position in drifting environments* (from definitions 1–3, background assumption 1, and assumptions 1, 2, 13–15, and 17–19).

$$\forall x, t_0, t_1, t_2 \{O(x) \wedge FG(x) \wedge \neg EN(x) \wedge [A(x, t_0) = 0] \wedge (\sigma > 0) \\ \wedge [A(x, t_1) \leq \sigma] \wedge [A(x, t_2) > \sigma] \rightarrow H(x, t_2) > H(x, t_1)\}.$$

In the case of robust position, three cases must be considered: (1) environmental change destroys alignment when advantage can be acquired from a robust position ($\sigma = \tau$), (2) environmental change destroys alignment more quickly than advantage can be acquired from a robust position ($\sigma < \tau$), and (3) advantage can be acquired from a robust position before environmental change destroys alignment ($\tau < \sigma$).

When $\sigma = \tau$, age dependence in the hazard for organizations with robust positions is monotonic negative:

THEOREM 8.—*Robust position without endowment when $\sigma = \tau$: If environmental drift destroys alignment exactly when advantage can be gained from occupancy of robust position ($\sigma = \tau$), then the hazard of mortality for an unendowed organization with a robust position decreases with age* (from definitions 1 and 2, background assumption 1, and assumptions 1, 2, 13–15, and 17–19).

$$\forall x, t_0, t_1, t_2 \{O(x) \wedge RB(x) \wedge \neg EN(x) \wedge [A(x, t_0) = 0] \wedge (\sigma > 0) \wedge (\tau > \sigma) \\ \wedge [A(x, t_1) \leq \sigma] \wedge [A(x, t_2) > \sigma] \rightarrow H(x, t_2) < H(x, t_1)\}.$$

Figure 5 shows the result for the other two cases, one in which $\sigma < \tau$ (on the left) and one with this inequality reversed (on the right). Note that, when drift produces environmental dissimilarity quickly relative to the speed of payoff to robust position, the hazard of mortality rises with age initially before declining (below the initial level)—there is a “liability of adolescence.”

THEOREM 9.—*Robust position without endowment when $\sigma < \tau$: If environmental drift destroys alignment before advantage can be gained from occupancy of robust position ($\sigma < \tau$), then the hazard of mortality for an unendowed organization with a robust position initially increases with age, then decreases with further aging and falls below the initial level* (from definitions 1, 2, and 4, background assumption 1, and assumptions 1, 2, 13–15, and 17–19).

$$\forall x, t_0, t_1, t_2, t_3 \{O(x) \wedge RB(x) \wedge \neg EN(x) \wedge [A(x, t_0) = 0] \wedge (\sigma > \tau) \wedge (\tau > 0) \\ \wedge [A(x, t_1) \leq \sigma] \wedge [A(x, t_2) > \sigma] \wedge [A(x, t_2) \leq \tau] \wedge [A(x, t_3) > \tau] \\ \rightarrow H(x, t_2) > H(x, t_1) > H(x, t_3)\}.$$

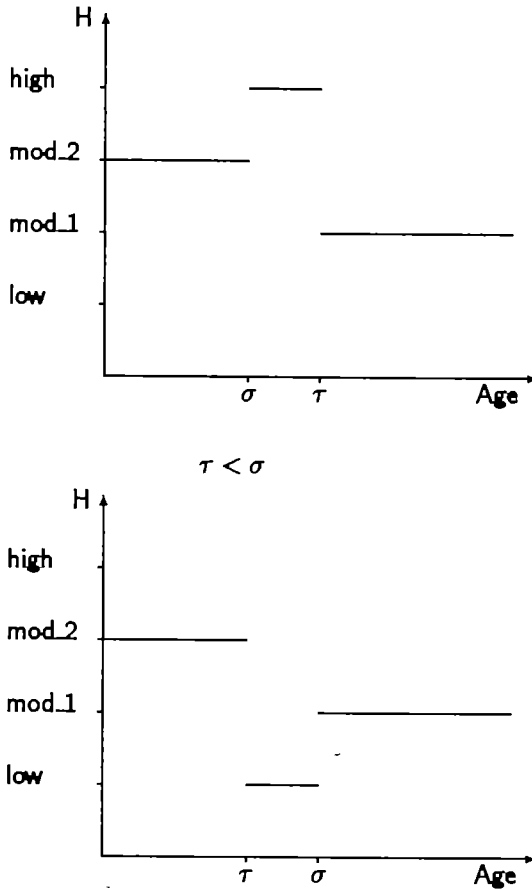


FIG. 5.—The implied pattern of age dependence in mortality hazards for unendowed organizations with robust positions for $\sigma < \tau$ and $\tau < \sigma$.

In the second case in figure 5 (on the right), when advantage can be gained from robust position before drift destroys alignment, then the mortality hazard peaks in youthful periods. The hazard falls with aging initially and then rises; but the hazard at later ages falls below the initial level of the hazard.

THEOREM 10.—*Robust position without endowment when $\sigma > \tau$: If advantage can be gained from occupancy of robust position before environmental drift destroys alignment ($\sigma > \tau$), then the hazard of mortality for an unendowed organization with a robust position initially decreases with age and then rises with further aging but remains below the initial level*

(from definitions 1, 2, and 4, background assumption 1, and assumptions 1, 2, 13–15, and 17–19).

$$\begin{aligned} \forall x, t_0, t_1, t_2, t_3 \{ & O(x) \wedge RB(x) \wedge \neg EN(x) \wedge [A(x, t_0) = 0] \wedge (\tau > 0) \wedge (\sigma > \tau) \\ & \wedge [A(x, t_1) \leq \tau] \wedge [A(x, t_2) > \tau] \wedge [A(x, t_2) \leq \sigma] \wedge [A(x, t_3) > \sigma] \\ & \rightarrow H(x, t_3) > H(x, t_1) > H(x, t_2) \}. \end{aligned}$$

Despite the difference in the “middle” period for these scenarios, the comparison between the early and late hazards is the same: the hazard for old organizations falls below the hazard for young ones. So this is one recipe for generating a result that parallels the new empirical findings on mortality rates in automobile industries (Hannan et al. 1998). Suppose that the form of positional advantage (fragile and robust) corresponds with position in the size distribution and that the organizations lack endowments. That is, assume that small organizations occupy fragile positions and that large ones occupy robust positions. Such an assumption does not appear unreasonable in the context of national automobile industries. Large firms have positional advantages in distribution (networks of dealers) that appear robust over many important kinds of changes in relevant environments. Such positions are rarely available to very small firms in these industries. Suppose that (1) position matters more than capability, (2) environmental and change erodes inert capabilities, (3) fragile position advantage can be gained at founding, but (4) advantage from robust position takes time to develop, then these theorems fit the result that age dependence in the hazard of mortality is positive for small (unendowed) organizations and negative for large (unendowed) ones.

The assumption that position dominates capability (assumption 19) matters in the foregoing argument for the case of organizations with robust positions. If we reversed this assumption, specifying that $mod_1 > mod_2$, then the “shape” of the relationship between age and the hazard would be the same as given in theorems 9 and 10. However, the comparison between the hazard near founding and the hazard at late ages would be different. Instead of the long-run hazard falling below the initial level (as in theorems 9 and 10), the long-run hazard would exceed the hazard in youth. That is, long-run age dependence would be *positive* for unendowed organizations with robust positions if capability dominates position in shaping life chances.

Case 2: organizations with endowments.—It is intuitively clear that adding endowments to the picture must create a pattern of long-run positive age dependence in the hazard. When organizations are assumed to begin with endowments, then the hazard of mortality for new organizations is *very low*, whether their positions are fragile or robust. Aging can

only increase the hazard, when the immunity ends and when drift eventually erodes alignment. The following theorem shows that this intuition is correct under the assumptions made in this section. (The intermediate steps needed to prove this theorem are reported in the appendix.)

THEOREM 11.—*The long-run hazard of mortality for an endowed organization with either a fragile or a robust position in a drifting environment exceeds the hazard near founding.*¹⁵

$$\begin{aligned} \forall x, t_0, t_1, t_2 \{ & O(x) \wedge [RB(x) \wedge FG(x)] \wedge EN(x) \wedge [A(x, t_0) = 0] \\ & \wedge (\sigma > 0) \wedge (\tau > 0) \wedge (\epsilon > 0) \wedge [A(x, t_1) \leq \min(\epsilon, \sigma, \tau)] \\ & \wedge [A(x, t_2) > \max(\epsilon, \sigma, \tau)] \rightarrow H(x, t_2) > H(x, t_1) \}. \end{aligned}$$

Proof.—Theorem A1 (in the appendix) states that the conclusion of this theorem hold for organizations with fragile positions. Lemmas A4 and A5 state that the conclusion holds for organizations with endowed, robust positions under the following circumstances: (1) $(\epsilon \geq \tau)$ and (2) $[\tau > \max(\epsilon, \sigma)] \wedge (\tau > \sigma)$. The only case that remains to be proven is for $\sigma > \tau > \epsilon$. According to the qualitative calculus (assumption 18), the hazard for the segment from founding to age ϵ equals *very_low* (because of the immunity conveyed by the endowment). In the rightmost segment, ages greater than σ , organizations have positional advantage ($\text{age} > \tau$) and lack alignment ($\text{age} > \sigma$). With these values of the functions, the hazard beyond age σ is *mod*₁, which exceeds *very_low* (by assumption 18). QED

V. DISCUSSION

The seemingly straightforward issue of age dependence in organizational mortality turns out on close inspection to be anything but simple. Inconsistency among theories and empirical findings abounds. Such inconsistency reflects the protean nature of the concept of organizational aging.

In one major line of work, aging means learning and solidifying external position. A contrary stream of research regards aging as increasing constraint on collective action, with the buildup of internal friction, precedent, and political deals combining to impede timely and reliable performance. The first set of formalizations presented here shows that these two lines of thinking can be cast as special cases of a general formulation in which aging can shape the cumulation of knowledge, internal friction,

¹⁵ This statement of the theorem uses comparisons involving the maximum of a pair of parameters. The theorem prover (OTTER), which was used to check proofs, does not include such expressions as basic “functors.” Therefore, lemmas and theorems with such expressions must be decomposed into sets of simpler ones that do not rely on the minimum and maximum operations.

and other constraints and the quality of external ties. The formalization also treats endowments as generating temporary (possibly partial) immunity from mortality. Adding endowments to the picture does not really change the substantive conclusion for the senescence argument. But it does matter in the liability-of-newness setup, changing the monotonic negative relation between age and the hazard to a nonmonotonic relation, with a liability of adolescence.

The general framework in which these three cases are embedded might serve as a basis for reconciling the conflicting arguments. Allowing the valences of the various effects of aging to vary systematically among different kinds of organizational populations and different contexts would help specify the substantive conditions under which each type of liability operates. Such an effort would also undoubtedly make sense of inconsistencies in empirical findings.

A radically different idea about aging lies at the heart of much of the theory and research on the determinants of organizational mortality. This is the idea that aging *per se* does not affect the hazard of mortality; instead, age tracks the fit between an organization and its environment. This image builds on ideas of imprinting and structural inertia. According to the imprinting hypothesis, organizations best match their external environments at the time of their founding. According to structural inertia theory, subsequent changes to the core of these imprinted features are hard and risky to the organizations' survival chances. If environments drift over time, then inert organizations fall further behind as they age—they become obsolete. Under this scenario, the hazard of mortality for old organizations exceeds the hazard for young ones.

The second set of formalizations offered here attempts to represent the logic of this general argument. In this set, age tracks the dissimilarity between an organization's current environment and its founding environment. This argument is specialized to the case of drifting environments, using the idea that there is a (common) age at which the environment becomes sufficiently different from the founding environment that it imposes fundamentally different demands on the organization. Organizations are assumed to be aligned well with their founding environments and not to be aligned with environments that are dissimilar to it. In the simpler version, an organization's capability is assumed to be higher when it is aligned with its environment. Thus, alignment unravels because environments drift and inert organizations continue to rely on their old structures and processes. Age tracks environmental drift in this formulation. Hence capability declines as organizations age, and the hazard of mortality rises with age.

A second version of this general argument refines the notion of position. It distinguishes two types of structural position: fragile and robust. The

difference between these types is that robust positions provide advantage over broader ranges of environmental variations. Fragile positions provide advantage immediately (at founding) but these advantages disappear when the environment becomes dissimilar. Robust positions do not provide advantage immediately; it takes time to build such positions. However, once established, robust positional advantages persist in all states of the environment. If organizations lack endowments, then age dependence is positive for the class of organizations with fragile positions and negative for those with robust positions. This difference in the pattern of age dependence by type of position might clarify the source of divergence in patterns of age dependence in recent the empirical research. For instance, suppose that organizational populations in drifting environments vary in the fractions of members with robust and fragile positions. If fragile positions predominate in a population, then simple age dependence will tend to be positive. But, if robust positions are the rule in some other populations, then simple age dependence will tend to be negative. Allowing endowments to affect the hazard of mortality does not change the result for organizations with fragile positions: age dependence is monotonically positive in drifting environments. However, endowment effects do complicate the picture for the case of robust positions. Under some of the specified conditions, the presence of an endowment produces the familiar "liability of adolescence." But, other conditions yield the opposite pattern, a "bathtub shape" relation between age and the hazard. The main result is that age dependence in the hazard of mortality is positive over the long run in drifting environments for endowed organizations with robust positions.

This article sought to clarify how the various stories "work" by interpreting them in FOL. The resulting formalizations are, of course, interpretations. Although they cannot reflect the full nuance of the natural language originals, they do place the common and divergent features of the contending theories in sharper relief.

However, these formalizations take only a first step in rationally reconstructing theories of age dependence. Some key issues still need to be addressed. Two seem to be especially pertinent. First, the review of theory and research on the issue indicated that age variations in organizational size plausibly play a role in generating patterns of age dependence. Obvious next steps would formalize these ideas in FOL and then seek to integrate the two formalizations. Second, there is an interesting tension between the obsolescence story and its formalizations here and the Hannan-Freeman (1984) theory of structural inertia. As formalized by Péli et al. (1997) the latter implies that the strength of structural inertia increases with age and also covaries with favorable selection chances (although inertia does not causally affect survival chances). The formalizations offered

here imply that obsolescence increases with age and that obsolescence varies inversely with survival chances. Yet, inertia is cited here as one of the causes of obsolescence. Resolving this tension, which requires constructing a more general formalization, lies beyond the scope of this article. But some of the directions that such an effort might take deserve mention.

Structural inertia theory does not emphasize environmental change. That is, this theory treats the consequences of the accumulating of reliability and accountability for reproducibility, inertia, and life chances without addressing issues concerning changes in the fit between an organization's routines and capabilities and the environment. In contrast, the formalizations of obsolescence presented here address environmental change directly but do not consider age-related changes in such survival-related properties as reliability and accountability of collective action. Suppose that elements of each account are correct. In particular, suppose that organizations gain reliability and accountability as they age and that routines and structures become obsolete as organizations age. Then as organizations age, they get better and better at doing things that convey less and less value for survival. In stable environments, getting better at doing old things aids survival more than doing the latest thing. In volatile environment, the situation is reversed. And there is a spectrum of intermediate possibilities. Efforts at formalizing these and related ideas would, by unifying several "theory fragments," take a valuable step toward building general, yet precise, theories of organization.

APPENDIX

Technical Background Assumptions

It simplifies equations to implement a trichotomy rule for the hazard (which can be interpreted to mean that "common sense" about inequalities applies):

BACKGROUND ASSUMPTION A1.—*A trichotomy rule holds for the hazard.*

$$\begin{aligned} \forall t_0, t \ (O(x) \rightarrow [[H(x, t) > H(x, t_0) \vee H(x, t) = H(x, t_0)] \\ \leftrightarrow \neg [H(x, t) < H(x, t_0)]]). \end{aligned}$$

A background assumption about inequalities is also employed throughout the analysis:

BACKGROUND ASSUMPTION A2

$$\forall x, y, z \ [(x > y) \wedge (y > z) \rightarrow (x > z)].$$

Intermediate Steps in Deriving Theorem 11

This appendix provides the intermediate steps needed to prove the theorem that states that endowed organizations in drifting environments experience long-run positive age dependence (theorem 11). There are several intermediate steps because many different cases need to be considered. In the case of fragile positions, the relevant cases are $\epsilon = \sigma$, $\epsilon > \sigma$, and $\epsilon < \sigma$.

LEMMA A1.—*Fragile position with endowment when $\epsilon = \sigma$: When endowments end when environmental drift destroys alignment ($\epsilon = \sigma$), then the hazard of mortality for an endowed organization with a fragile position remains constant during the period of immunity and jumps when immunity ends* (from definitions 2–3, background assumption 1, and assumptions 1, 2, 13–15, and 17–18).

$$\forall x, t_0, t_1, t_2 \{O(x) \wedge FG(x) \wedge EN(x) \wedge [A(x, t_0) = 0] \wedge (\sigma > 0) \wedge (\epsilon > \sigma) \\ \wedge [A(x, t_1) \leq \epsilon] \wedge [A(x, t_2) > \epsilon] \rightarrow H(x, t_2) > H(x, t_1)\}.$$

LEMMA A2.—*Fragile position with endowment when $\epsilon < \sigma$: When endowments end before environmental drift destroys alignment ($\epsilon < \sigma$), then the hazard of mortality for an endowed organization with a fragile position stays constant during the period of immunity, jumps when immunity ends and jumps further when drift eliminates alignment* (from definitions 1–3, background assumption 1, and assumptions 1, 2, 13–15, and 17–19).

$$\forall x, t_0, t_1, t_2, t_3 \{O(x) \wedge FG(x) \wedge EN(x) \wedge [A(x, t_0) = 0] \wedge (\epsilon > 0) \wedge (\sigma > \epsilon) \\ \wedge [A(x, t_1) \leq \epsilon] \wedge [A(x, t_2) > \epsilon] \wedge [A(x, t_2) \leq \sigma] \wedge [A(x, t_3) > \sigma] \\ \rightarrow H(x, t_3) > H(x, t_2) > H(x, t_1)\}.$$

LEMMA A3.—*Fragile position with endowment when $\epsilon > \sigma$: When endowments end after environmental drift destroys alignment ($\epsilon > \sigma$), then the hazard of mortality for an endowed organization with a fragile position in a drifting environment is constant during the period of immunity and jumps when immunity ends* (from definitions 1–3, background assumption 1, and assumptions 1, 2, 13–15, and 17–19).

$$\forall x, t_0, t_1, t_2, t_3 \{O(x) \wedge FG(x) \wedge EN(x) \wedge [A(x, t_0) = 0] \wedge (\sigma > 0) \wedge (\epsilon > \sigma) \\ \wedge [A(x, t_1) \leq \sigma] \wedge [A(x, t_2) > \sigma] \wedge [A(x, t_2) \leq \epsilon] \wedge [A(x, t_3) > \epsilon] \\ \rightarrow H(x, t_3) > H(x, t_2) > H(x, t_1)\}.$$

THEOREM A1.—*The hazard of mortality for an endowed organization with a fragile position rises with age after the period of immunity* (from lemmas A1–A3).

$$\forall x, t_0, t_1, t_2 \{O(x) \wedge FG(x) \wedge EN(x) \wedge [A(x, t_0) = 0] \wedge (\sigma > 0) \wedge (\epsilon > 0) \\ \wedge [A(x, t_1) \leq \epsilon] \wedge [A(x, t_2) > \epsilon] \rightarrow H(x, t_2) > H(x, t_1)\}.$$

Now consider organizations with robust positions. Now there are 16 cases to be considered, because three parameters are relevant: ϵ , σ , and τ . Examination of these cases reveals that they can be collapsed into three sets. First, endowments can persist until positional advantage accrues to robust position ($\epsilon \geq \tau$). In this case, the hazard is constant at the level *very low* until the immunity ends. If drift has already eliminated alignment, then the hazard jumps immediately. Otherwise it jumps when drift subsequently eliminates alignment. Thus this case parallels the one for organizations with robust position but without endowments (theorem 8): age dependence in the hazard is monotonically positive.

LEMMA A4.—*Robust position with endowment when $\epsilon \geq \tau$: The hazard of mortality for an endowed organization with a robust position rises monotonically with age when $\epsilon \geq \tau$* (from definitions 1, 2, and 4, background assumption 1, and assumptions 1, 2, 13–15, and 17–19).

$$\forall x, t_0, t_1, t_2 \{O(x) \wedge RB(x) \wedge EN(x) \wedge [A(x, t_0) = 0] \wedge (\sigma > 0) \\ \wedge (\epsilon > 0) \wedge (\epsilon \geq \tau) \wedge [A(x, t_1) \leq \epsilon] \wedge [A(x, t_2) > \epsilon] \\ \wedge [A(x, t_2) > \sigma] \rightarrow H(x, t_2) > H(x, t_1)\}.$$

The second important case is one in which there is a nonzero interval between the ending of the endowment and the onset of position advantage from robust position ($\tau > \epsilon$) and positional advantage from robust position occurs after the loss of alignment due to drift ($\tau > \sigma$).

LEMMA A5.—*Robust position with endowment when $\tau > \max(\epsilon, \sigma)$: The hazard of mortality for an endowed organization with a robust position rises and then falls with age with the long-run level of the hazard exceeding the initial level when $\tau > \epsilon$ and $\tau > \max(\epsilon, \sigma)$* (from definitions 1, 2, and 4, background assumption 1, and assumptions 1, 2, 13–15, and 17–19).

$$\forall x, t_0, t_1, t_2, t_3 \{O(x) \wedge RB(x) \wedge EN(x) \wedge [A(x, t_0) = 0] \wedge (\sigma > 0) \\ \wedge (\epsilon > 0) \wedge (\tau > \epsilon) \wedge (\tau > \sigma) \wedge [A(x, t_1) \leq \epsilon] \\ \wedge [A(x, t_2) > \epsilon] \wedge [A(x, t_2) \leq \tau] \wedge [A(x, t_3) > \tau] \\ \rightarrow H(x, t_2) > H(x, t_3) > H(x, t_1)\}.$$

The third case is one in which $\sigma > \tau > \epsilon$, which is discussed in the proof of theorem 11.

REFERENCES

- Barron, David N., Elizabeth West, and Michael T. Hannan. 1994. "A Time to Grow and a Time to Die: Growth and Mortality of Credit Unions in New York, 1914–1990." *American Journal of Sociology* 100:381–421.
- Bruderl, Josef, and Rudolph Schüssler. 1990. "Organizational Mortality: The Liabilities of Newness and Adolescence." *Administrative Science Quarterly* 35:530–37.
- Bruderl, Josef, Peter Preisendörfer, and Rolf Ziegler. 1996. *Der Erfolg neugegründeter Betriebe*. Berlin: Duncker & Humbolt.
- Burt, Ronald. 1992. *Structural Holes: The Social Structure of Competition*. Cambridge, Mass.: Harvard University Press.
- Carroll, Glenn R. 1983. "A Stochastic Model of Organizational Mortality: Review and Reanalysis." *Social Science Research* 12:303–29.
- Carroll, Glenn R., and Michael T. Hannan. 1989. "Density Delay in the Evolution of Organizational Populations: A Model and Five Empirical Tests." *Administrative Science Quarterly* 34:411–30.
- Fichman, Mark, and Daniel A. Levinthal. 1991. "Honeymoons and the Liability of Adolescence: A New Perspective on Duration Dependence in Social and Organizational Relationships." *Academy of Management Review* 16:442–68.
- Freeman, John, Glenn R. Carroll, and Michael T. Hannan. 1983. "The Liability of Newness: Age Dependence in Organizational Death Rates." *American Sociological Review* 48:692–710.
- Freeman, John, and Michael T. Hannan. 1983. "Niche Width and the Dynamics of Organizational Populations." *American Journal of Sociology* 88:1116–45.
- Hannan, Michael T. 1996. "Selection Mechanisms: Organizations and Other Corporate Actors." Paper presented at the Conference on Social Mechanisms, Stockholm.
- Hannan, Michael T., M. Diane Burton, and James N. Baron. 1996. "Inertia and Change in the Early Years: Employment Relations in Young, High-Technology Firms." *Industrial and Corporate Change* 5:503–36.
- Hannan, Michael T., and John Freeman. 1977. "The Population Ecology of Organizations." *American Journal of Sociology* 82:929–64.
- . 1984. "Structural Inertia and Organizational Change." *American Sociological Review* 49:149–64.
- Hannan, Michael T., Glenn R. Carroll, Stanislav Dobrev, and Joon Han. 1998. "Age Dependence in Organizational Mortality Revisited, Part I: European and American Automobile Industries, 1885–1981." *European Sociological Review* 14, in press.
- Kreps, David M. 1996. "Markets and Hierarchies and (Mathematical) Economic Theory." *Industrial and Corporate Change* 5:561–96.
- Levinthal, Daniel. 1991. "Random Walks and Organizational Mortality." *Administrative Science Quarterly* 36:397–420.
- March, James G. 1988. *Decisions in Organizations*. New York: Blackwell.
- . 1991. "Exploration and Exploitation in Organizational Learning." *Organization Science* 2:71–87.
- McCune, William W. 1994. *Otter 3.0 Reference Manual and Guide*, rev ANL-94/6. Argonne, Ill.: Argonne National Laboratory.
- Nelson, Richard R., and Sidney G. Winter. 1982. *The Evolutionary Theory of Economic Change*. Cambridge, Mass.: Belknap Press.
- Péll, Gábor. 1997. "The Niche Hikers Guide to Population Ecology: A Reconstruction of Niche Theory Using Logic." *Sociological Methodology* 27:1–46.
- Péll, Gábor, Jeroen Bruggeman, Michael Maus, and Breannán Ó Nualláin. 1994. "A Logical Approach to Organizational Ecology: Formalizing the Inertia-Fragment in First-Order Logic." *American Sociological Review* 59:571–93.

- Péli, Gábor, and Michael Masuch. 1997. "The Logic of Propagation Strategies: Axiomatizing a Fragment of Organizational Ecology in First-Order Logic." *Organization Science* 3:310–31.
- Péli, Gábor, László Pólos, and Michael T. Hannan. 1997. "Back to Inertia: New Logical Formalizations." Technical report. University of Groningen, Faculty of Management and Organization.
- Pólos, László, and Jaap Kamps. 1997. "Towards a Formal Theory of Organizations in Action." Manuscript. University of Amsterdam, Applied Logic Laboratory.
- Podolny, Joel. 1993. "A Status-Based Model of Market Competition." *American Journal of Sociology* 98:829–72.
- Podolny, Joel M., Toby E. Stuart, and Michael T. Hannan. 1996. "Networks, Knowledge, and Niches: Competition in the Worldwide Semiconductor Industry, 1984–1991." *American Journal of Sociology* 102:659–89.
- Stinchcombe, Arthur L. 1965. "Social Structure and Organizations." Pp. 142–93 in *Handbook of Organizations*, edited by J. G. March. Chicago: Rand McNally.
- Thompson, James D. 1967. *Organizations in Action*. New York: McGraw-Hill.

Uncertainty, Trust, and Commitment Formation in the United States and Japan¹

Toshio Yamagishi
Hokkaido University

Karen S. Cook
Duke University

Motoki Watabe
University of California, Los Angeles

A theory of trust proposed by Yamagishi and Yamagishi provides the basis for the prediction that (1) social uncertainty promotes commitment formation between particular partners and (2) high trusters tend to form committed relations less frequently than would low trusters when facing social uncertainty. These predictions receive support in two experiments conducted in the United States and Japan. The findings provide empirical support for the theory of trust that emphasizes the role of general trust (trust in others in general) in emancipating people from the confines of safe, but closed relationships. The results also offer a theoretical explanation for what have been viewed in the past as cultural differences.

INTRODUCTION

Intense family ties prevent trust from developing beyond the confines of the family. This is the central thesis of Fukuyama's (1995) book on trust. From a similar theoretical perspective, Yamagishi and his associates (Kiyonari and Yamagishi 1996; Yamagishi and Yamagishi 1989; Yamagishi

¹ The research presented in this article was supported by scientific research grants from the Institute of Nuclear Safety Systems and the Japanese Ministry of Education. We would like to thank J. Misumi of the Institute of Nuclear Safety Systems for his support and encouragement of our research project. We are also grateful to Nahoko Hayashi, Nobuhito Jin, Nobuyuki Takahashi, Yoriko Watanabe, Motoko Kosugi, Masanori Takezawa, Shawn Donnelly, and Bob Duniway for their assistance in conducting experiments. Please address correspondence to Toshio Yamagishi, Faculty of Letters, Hokkaido University, N10 W7 Kita-ku, Sapporo, Japan 060. E-mail: Toshiolet.hines.hokudai.ac.jp

1995c, 1995d, 1996, in press; Yamagishi and Komiyama 1995; Yamagishi et al. 1996; Yamagishi and Yamagishi 1994, in press; Yamagishi et al. 1995) have conducted research that confirms the proposition that intense group ties, often observed in collectivist cultures, prevent trust from developing beyond group boundaries. The proposition—strong family and group ties confine “trust” to within the boundaries of the family and the group—advanced by these and other scholars (e.g., Gellner 1988; Hawthorn 1988) makes intuitive sense on the one hand, but at the same time it contains counterintuitive implications. In particular, it implies that societies characterized by a prevalence of strong social ties (such as Japanese society) produce less trust among their members than do societies in which social and interpersonal ties are weaker. This implication is counterintuitive since it is commonly believed that a society characterized by strong social and interpersonal relations, such as Japanese society, is also characterized by a high level of trust among its members.

Despite this common view and consistent with the implication of the above proposition, cross-societal research comparing the United States and Japan has repeatedly demonstrated that the level of general trust is much higher in American society than in Japanese society. First, based on the above proposition, Yamagishi and Yamagishi (1994) derived the prediction that the average level of general trust (i.e., trust in others in general) would be higher among Americans than among the Japanese and confirmed this prediction in results obtained from a cross-societal questionnaire survey. Although this survey did not use nationally representative samples, another, more systematic study conducted by the Institute of Statistical Mathematics (Hayashi et al. 1982) using representative national samples reports a similar cross-national difference. According to this study, 47% of the American sample ($N = 1,571$) responded “People can be trusted” to the question, “Do you think you can put your trust in most people, or do you think it’s always best to be on your guard?” In contrast, only 26% of the Japanese sample ($N = 2,032$) gave the same response. Similarly, 47% of the American sample, compared to 19% of the Japanese sample, answered that people try to be helpful when asked, “Would you say that most of the time people try to be helpful, or that they are mostly just looking out for themselves?”

Building on the proposition that strong and stable social relations (such as family ties and group ties) promote a sense of security within such relations but endanger trust that extends beyond these relations, Yamagishi and Yamagishi (1994) provided an explanation for the counterintuitive, but consistent, finding that Americans are more “trusting” than the Japanese. The purpose of this study is to provide further empirical support for the theory of trust proposed by Yamagishi and his associates. After

describing the theory we present findings from two cross-societal experiments conducted in the United States (Seattle) and Japan (Sapporo).

INSTITUTIONAL VERSUS INDIVIDUALISTIC VIEWS OF CULTURE

The theory of trust developed by Yamagishi and his associates is primarily based on the contrast between institutional and individualistic views of culture. Yamagishi (1988a) compared the tendencies of American and Japanese subjects to desert to a group that contained free riders. The common view of the cultures in American and Japanese societies—the former characterized by individualism and the latter by collectivism—would imply that Japanese subjects would have a stronger preference to stay in the group than Americans would, given the existence of free riding on the part of other group members. Contrary to this simplistic view of American and Japanese cultures, the results of this experiment indicated that the tendency to remain in the group was much stronger among the American subjects than it was among the Japanese. More specifically, American subjects exited the group on only slightly more than one out of 20 trials when they incurred monetary costs (i.e., in the high-exit-cost condition) for leaving, whereas Japanese subjects exited the group much more frequently (on about eight out of 20 trials), despite the fact that they earned much less money by doing so. This difference in exit rates did not exist between American and Japanese subjects when the exit costs were low (i.e., in the low-exit-cost condition). These results indicate that both American and Japanese subjects disliked staying in a group with free riders, as evidenced by the high level of exit responses among both groups in the low-cost condition. However, American subjects stayed in the group when the exit response was costly, whereas Japanese subjects were willing to pay extra for the exit option. These results would be difficult to explain in terms of the individualistic view of culture. If Japanese people are collectivist in the sense that they prefer to be part of the group rather than independent from the group, and if American people are individualistic in the sense that they value independence from the group, then American subjects should have exited more often than did the Japanese subjects in this experiment regardless of cost.

This seemingly counterintuitive pattern of findings was predicted by Yamagishi (1988a) based on what he calls the “institutional view” of culture. According to this view, the Japanese often “prefer” to belong to groups and place group interests above their own individual interests not because they intrinsically like to do so, but because it is in their own long-term interest. In the context of free riding in a collective work group, Japanese society has developed systems of mutual monitoring and sanc-

tioning to curtail free riding (Hechter and Kanazawa 1993), and these solutions work for the group insofar as such a "collective solution" to the free rider problem is in place. However, in groups artificially created in the laboratory without opportunities for face-to-face interaction (i.e., in the groups used in Yamagishi's experiment) such collective solutions as informal mutual monitoring and sanctioning do not exist. In such a situation, Japanese people tend to prefer not to stay in the group, as indicated by the experimental results. A collectivist culture did not exist among the Japanese subjects who participated in this experiment because the artificially created groups lacked opportunities for the mutual monitoring and sanctioning of each other's behavior.

Another experiment by Yamagishi (1988b) provides an even clearer contrast between the individualistic view of culture and this institutional view of culture. In a cross-societal experiment between the United States and Japan comparing cooperative tendencies in social dilemmas, subjects in four-person groups were each given 50¢ (U.S.) or ¥100 (Japan)¹ and were asked how much of the money they wanted to provide for the welfare of the other group members. They could contribute any amount between zero and 50¢ or ¥100. The amount provided by the subject was doubled in value and then equally allocated among the other three members. Thus, if all four contributed the full amount, each earned \$1 or ¥200. However, the subject could have earned more (\$1.50/¥300) if he or she had not contributed any while the other three fully contributed. The decision was repeated 12 times (Japan) or 16 times (U.S.), and thus the pure free rider could earn \$1.50 times 16 (\$24), or ¥300 times 12 (¥3,600). These amounts were smaller than those they would have obtained if all four kept contributing in full (\$16 or ¥2,400). Thus, not contributing at all was the dominant or "rational" choice. However, if everyone had chosen this dominant response each would have earned only 50¢ or ¥100 per trial and \$8 or ¥1,200 in total. This situation represents a social dilemma in which (1) each member has a choice between C (cooperation, contributing the full amount) and D (defection, contributing nothing), (2) the choice of D produces a better outcome for a member than the choice of C regardless of the choices of the other members, and yet (3) the individual outcome when all members choose D is worse than the individual outcome when all choose C (see Dawes [1980], Messick and Brewer [1983], and Yamagishi [1995b], for reviews of the social dilemma literature).

The individualistic view of culture—that Japanese individuals value group interests over individual interests more so than Americans do—predicts that Japanese subjects are more likely to cooperate in this experiment than are American subjects. On the other hand, the institutional

¹ At the time of the experiment, \$1.00 was worth about ¥200.

view of culture—that Japanese individuals cooperate with the group because there exists a system of formal and informal mutual monitoring and sanctioning—predicts that Japanese subjects will cooperate less than American subjects will in such an experimental social dilemma. This is because the Japanese are so accustomed to an environment in which collective systems of mutual sanctioning guarantee mutual cooperation in the group that they would feel insecure in a social environment lacking such a system of mutual monitoring and sanctioning. Using this reasoning, Yamagishi (1988*b*) predicted that Japanese subjects would have a lower level of trust in strangers and would cooperate less in social dilemmas involving strangers compared to the behavior of American subjects. This prediction also received clear empirical support. Contrary to the individualistic view of culture, and consistent with the institutional view, American subjects in this experiment cooperated more often (contributing 56.2% of 50¢ per trial) than did Japanese subjects (who contributed 44.4% of ¥100 per trial). Furthermore, the questionnaire data showed that the American subjects were more trusting than the Japanese were.³ That is, in a group consisting of strangers, lacking opportunities for mutual sanctioning, American subjects voluntarily cooperated and contributed more to the welfare of the group than the Japanese subjects did. On the other hand, in another condition in which subjects were given opportunities to sanction each other, the Japanese subjects became much more cooperative. With the addition of opportunities for mutual sanctioning, Japanese subjects' average cooperation level increased from 44.4% to 74.6%, an improvement of 30 percentage points. In contrast, the sanctioning opportunities did not have as strong an effect on the Americans. The sanctioning opportunities improved the average cooperation level of the American subjects' from 56.2% to 75.5%, an improvement of 19 percentage points. It was thus shown that the "collectivist" behavior—cooperation for the welfare of the group as a whole—among the Japanese subjects was maintained to a large extent by the system of mutual monitoring and sanctioning, rather than by their presumed "value system" according to which each individual places the group's welfare above his or her own self-interest.

A THEORY OF TRUST

Yamagishi and his associates have developed the "emancipation" theory of trust as a theoretical extension of the institutional view of culture pro-

³ Earlier studies by Yamagishi and his associates (Yamagishi 1988*b*, Yamagishi and Yamagishi 1994) also report a similar difference in the level of general trust between American and Japanese respondents.

posed in their cross-societal experiments.⁴ According to this theory, general trust (trust in others in general) and commitment formation are considered alternative solutions to the problems caused by social uncertainty. Social uncertainty is ubiquitous in human society. Whenever we interact with others we face the problem of social uncertainty. We engage in social interactions with others to improve our own welfare, material or psychological; however, in interacting with others we make ourselves vulnerable. We seek to improve our welfare while taking the risk of incurring costs. We use the term "social uncertainty" to refer to the risk of being exploited in social interactions. That is, social uncertainty is defined as existing for an actor when (1) his or her interaction partner has an incentive to act in a way that imposes costs (or harm) on the actor and (2) the actor does not have enough information to predict if the partner will in fact act in such a way.

Kollock (1994) provides a good example of how social uncertainty promotes commitment between particular exchange partners.⁵ The examples he uses are rice and rubber markets in Southeast Asia. The quality of rice is immediately apparent upon simple inspection. The buyer has little risk of being cheated on the quality of rice he buys, and thus he faces a low level of social uncertainty. In contrast, the quality of raw rubber is hard to judge; its quality can be known only after it has been processed. Cheating on quality is easier, and the consequence of being cheated in this situation is extremely serious. In other words, the buyer of raw rubber thus faces a high level of social uncertainty. This difference in social uncertainty concerning the quality of rice and rubber, Kollock argues, explains the observed difference in the dominant form of trade. Rice is usually traded at open markets between strangers, whereas rubber is often traded between a particular producer and a broker who have formed a long-term relationship, often extending over several generations. A high level of social uncertainty concerning the quality of rubber is the determining factor for the development of such committed relations between rubber producers and brokers (see also Akerlof 1970). The experiment Kollock conducted is a laboratory version of rice and rubber trades. In one condition (the high uncertainty condition), sellers could lie to their

⁴ We call these experiments "cross-societal" rather than "cross-cultural" since their purpose was to show that cultural differences can be often be explained in terms of other, more tangible theoretical variables.

⁵ The term "commitment" in this article is used in a strictly behavioral manner. One is defined as committed to a relationship to the degree that he or she forgoes better alternatives. Mutual attraction, liking, and loyalty may emerge in such a committed relation, and when they do, they will certainly strengthen the commitment. Nevertheless, such psychological factors, however strongly related to commitment, are not commitment itself in this sense.

potential buyers concerning the quality of the product they sold. In the other condition (the low uncertainty condition), sellers could not lie. The results of Kollock's experiment demonstrated that commitment formation between a particular seller and a particular buyer occurred more frequently in the high uncertainty condition than in the low uncertainty condition.

Commitment formation as a solution to problems of social uncertainty, however, has its own problems. While reducing the risk of being duped in interacting with unfamiliar people,⁶ it limits the actor's choices for exploring better opportunities that might exist outside the current relationship. Using terminology from economics, commitment formation reduces transaction costs, on the one hand, but imposes opportunity costs, on the other. In forming a commitment with a particular partner, one obtains security (i.e., a reduction in social uncertainty) in exchange for opportunities. Commitment formation is an efficient means for reducing uncertainty in a situation in which outside opportunities are limited (i.e., when the general level of opportunity costs is low). On the other hand, such commitments become a liability rather than an asset as people face more and better opportunities outside their current, mutually committed relationships (i.e., when the general level of opportunity costs in the environment is high). General trust (or trust in people in general) provides a springboard for people who have been "confined" to committed relationships to move out into the larger world of opportunities.⁷ In this way general trust

⁶ According to Yamagishi and Yamagishi (1994), social uncertainty is reduced in committed relations for the following reasons. First, committed partners accumulate information about each other. Second, "hostage-taking" behaviors, which range from the formation of mutual emotional attachments to the establishment of relation-specific assets (Helper and Levine 1992), develop in mutually committed relations; these hostage-taking behaviors provide deterrence against unilateral defection (Shapiro, Sheppard, and Cheraskin 1992). Finally, strategies such as tit-for-tat can be used to control each other's behavior (cf. Axelrod 1984).

⁷ Yamagishi and Yamagishi (1994) mention four reasons for the difficulty people have in leaving a committed relationship even when it becomes a liability. First, committed people are by definition those who stay with the current relations despite outside opportunities (see n. 3 above). They may eventually leave the committed relationship, but there typically is a substantial time lag. Secondly, the mutual attraction and loyalty that have developed through the relationship keep partners in the relationship. Third, a temporary better offer from outsiders would not be sufficient for the one who has invested in relation-specific assets (see n. 4 above) to leave the current relationship. Social and psychological assets, such as the warm memory of a pleasant past and mutual understanding, may be considered relation-specific assets that keep people in these relationships. Finally, commitment to a particular partner often reduces the level of trust in "outsiders" (see Kiyonari and Yamagishi [1996] for experimental support), creating a vicious cycle of distrust of outsiders; those who do not trust "outsiders" tend to stay in committed relationships, and because they avoid "outsiders" they become even less trusting of "outsiders."

emancipates people from the confines and security of stable, committed relationships.⁴

In Japanese society the stability of interorganizational as well as interpersonal relationships makes exploitative, short-term profit-maximizing behavior less profitable than in American society. The one who deserts a relationship for quick profit will have a harder time in a society in which other relationships are mainly closed to outsiders. In other words, the stable nature of social and organizational relationships reduces social uncertainty and provides security inside of such relationships. This sense of security is what is often considered "trust" when characterizing the Japanese scene. However, once such sociorelational bases of security are removed, the Japanese may feel more insecure than Americans do. When facing strangers outside of established relationships, the Japanese are more distrustful of others in general. In short, in contrast to Americans, Japanese feel a greater sense of security within established and stable relationships but are more distrustful of people outside of the context of such relationships.

In the experiments presented below, we conduct empirical tests of critical aspects of the theory of trust briefly described above. The first experiment tests the proposition that social uncertainty facilitates commitment formation. In the second experiment, we test another critical proposition that trust emancipates people from committed relationships. These two experiments were conducted in the United States and in Japan in order to demonstrate that the theory applies to both societies and helps to explain what have previously been viewed as strictly cultural differences.

⁴ We do not discuss here the mechanism by which people living in such an environment come to acquire a high level of general trust. Being fully aware of the potentially quasi-functional nature of our argument without such a mechanism, we are proposing elsewhere how investment in cognitive resources provides such a mechanism (Yamagishi 1995a, 1996), according to which general trust is a by-product of cognitive investments. That is, those who have dedicated a significant proportion of their cognitive resources to developing skills to discern trustworthy from untrustworthy people can afford to maintain high default expectations of others' levels of trustworthiness. By maintaining a high level of general trust, they enjoy the advantage of being able to explore fully those opportunities that lie outside their established relationships. At the same time, they can quickly pull out of potentially harmful or costly relationships at the first sign of risk. Those who have not made such cognitive investments are slow in detecting the signs of untrustworthiness in their partners and thus are less likely to explore potentially fruitful, but risky relationships. Consistent with this model, several experimental studies (Kikuchi, Watanabe, and Yamagishi 1995; Kosugi 1996; Kosugi and Yamagishi 1995) have shown that compared to low trusters, high trusters are more sensitive to information potentially revealing other people's trustworthiness and are more accurate in discerning trustworthy people from those who are untrustworthy.

EXPERIMENT 1

The first study examines the key proposition tested by Kollock (1994) in his investigation of differences in situations such as those reflected in the comparison of rice and rubber markets. Kollock's experiment simulated transactions between buyers and sellers. In one condition (the uncertainty-quality condition), the true quality of the goods sold by sellers was known only to the seller. The seller announced the quality of the goods he or she offered to sell to a potential buyer, but only the seller knew the truth. Buyers in this condition could be deceived concerning the quality of the goods they were buying. In the certain-quality condition, the quality of all goods offered for sale were transparent to everyone, including both the sellers and the buyers. We also simulated trading practices between buyers and sellers. However, our experiment differed from Kollock's in three respects. First, only one person (a buyer) was the actual human subject in our study. The rest were "simulated actors" or programmed responses by a computer. This change provided us with a chance to test the predictions under a more controlled environment. Second, we developed a more powerful manipulation of social uncertainty. In addition to allowing sellers (through simulated responses) not to reveal the true quality of their goods, we also gave them the chance to extort money from the buyer. The potentially negative consequences of buying from the wrong seller in this experiment are thus much more serious than in Kollock's original experiment. Finally, we replicated the same experiment in the United States and in Japan to investigate the generality of the proposition across two different cultures. Although the literature on Japanese society and Japanese business emphasizes the stable and long-term nature of interpersonal and interorganizational relations in Japan, we predict that the level of commitment formation would be the same among Americans and Japanese when they face the same level of social uncertainty and when they are matched on their levels of trust. Specifically, we tested the following two hypotheses derived from our theory of trust.

HYPOTHESIS 1.—*Social uncertainty will facilitate commitment formation between particular partners.*

This hypothesis was originally developed by Cook and Emerson (1978) and was successfully tested by Kollock (1994). According to Kollock, when faced with a high level of social uncertainty or "a situation in which one can be taken advantage of, the natural response is to restrict one's transactions to those who have shown themselves to be trustworthy (i.e., becoming committed to particular exchange partners)" (1994, p. 318). Furthermore, construction of a mutually committed relationship transforms the "one-shot" interactions into "iterated" interactions, using game theoretic terms, and thus makes cooperative choices individually more profitable

than noncooperative choices as the prisoner's dilemma literature (cf. Axelrod 1984) indicates. The expectation of future benefits removes uncertainty (i.e., the chance of being exploited by the partner) from a mutually committed, long-term relationship. In other words, the benefit expected in the future provides assurance of the partner's willingness to act in a benign manner. Finally, the literature on prisoner's dilemma networks (Hayashi 1993, 1995; Hayashi, Jin, and Yamagishi 1993; Jin, Hayashi, and Shinotsuka 1993; Schuessler 1989; Vanberg and Congleton 1992; Yamagishi and Hayashi 1996; Yamagishi, Hayashi, and Jin 1994) provides further experimental and simulation evidence that people facing social uncertainty tend to form mutually committed relationships. The situation studied in this literature is a "prisoner's dilemma network" in which people select partners to play prisoner's dilemma games with (i.e. interaction situations characterized by a high level of social uncertainty). It has been established in this literature that one of the most effective strategies in prisoner's dilemma networks is the tit-for-tat strategy (Axelrod 1984; Hayashi 1993, 1995), according to which a player keeps cooperating with a cooperative partner and deserts the partner as soon as he or she fails to cooperate. The player who has adopted this strategy becomes committed to a cooperative partner, and if the partner has also adopted the same strategy they form a mutually committed relation. In all of these scenarios, it is social uncertainty (i.e., the risk of being exploited in social interactions) that drives people to form mutually committed, "safe" relationships.

HYPOTHESIS 2.—The level of commitment formation among Japanese subjects will not be different from that among American subjects when the level of social uncertainty and trust are made equivalent in the two groups.

The theory of trust proposed by Yamagishi and Yamagishi (1994) posits two major determinants of commitment formation. The first is the level of social uncertainty discussed above. The second is the level of the actor's general trust. People with a low level of general trust would prefer not to deal with anyone outside their established relations and thus tend to maintain committed relations at a higher level than do those having a higher level of general trust. The prediction concerning the effect of general trust, however, will not be tested in the first experiment; it will be tested in the second experiment. Instead, we matched Japanese and American subjects with respect to their levels of trust. Once these two factors (level of trust and amount of social uncertainty) are controlled, we have no theoretical reason to expect a cross-societal difference in the level of commitment formation, and thus we predict that the Japanese participants are not more likely to form committed relations than are the Americans.

METHOD

The experiment simulated transactions between buyers and sellers. Each experimental group consisted of several buyers and several sellers, but the subject (who was "randomly" assigned to the role of a buyer) was the only human subject in the group. The rest, including all sellers and buyers other than the subject, were actually programmed responses from the computer.

Subjects

Subjects were recruited from various classes at the University of Washington and at Hokkaido University. At the time of recruitment, at least a few weeks before the experiment, subjects were asked to fill out an eight-item trust scale developed by Yamagishi (1986, 1988*b*, 1988*c*, 1992). Potential subjects recruited into the subject pool were divided into the high-trust group and the low-trust group using a median split.⁹ Equal numbers of high trusters and low trusters (50 each) were selected from the subject pool in each country. Although previous studies (Yamagishi 1988*b*; Yamagishi and Yamagishi 1989, 1994) consistently produced the result that Americans have a higher level of general trust than the Japanese, this cross-cultural difference in the level of general trust was removed as a possible factor by matching the subjects who participated in this experiment on trust levels. We originally planned to use subjects' levels of general trust as an additional factor, hoping to test the proposition that low trusters are more likely to form committed relations than are high trusters, but the data for the American subjects' individual trust scores were not recorded in such a way as to allow individual \times individual linkage of scores to specific subjects. (This file was destroyed in data shipment.) The scores were available by condition because they were used to assign individuals to conditions, but they were not available by individual for subsequent analyses. Therefore, we know that the American and Japanese subjects are matched with regard to their trust scores by condition, but we cannot use trust as an independent factor in the analysis.

Five or six subjects were scheduled for each experimental session. When fewer than five subjects showed up for a session, confederates were brought into the laboratory to create the atmosphere that at least five or six subjects were participating. Although subjects were led to a private compartment as soon as they arrived at the laboratory and did not have

⁹ Subjects whose trust scores were 28 or higher were classified as high trusters, and low trusters had trust scores of 20 or lower. The median trust score was 25.

a chance to meet other participants, the compartments were not well insulated and thus subjects could hear the experimenter talking to other participants.

Design

The manipulation of social uncertainty (low vs. high) was crossed with the subjects' nationality (United States vs. Japanese).

General Procedures

The experiment was conducted in 1993 in Japan and in 1994 in the United States. In each country, the experiment took place in a comparably designed social psychology laboratory complex consisting of several compartments for the subjects and a control room. Subjects were provided with relevant information displayed on a computer located in front of them in each compartment, and their responses were entered into the computer through the keyboard. The subjects' computers were networked to the host computer located in the control room. Subjects had no opportunity to meet the other participants in person before, during, or after the experiment.

Upon arrival at the laboratory complex, each subject was led to a compartment and was handed the first set of instructions. The instructions told them that participants play the role of either a buyer or a seller who repeatedly conduct transactions. The last part of the instructions asked the subject to hit a key on a computer keyboard to start "random assignment" of the role. While subjects were told that the role assignments would be random, only the role of "buyer" was assigned to them. The experimenter then came into the compartment, gave further instructions concerning how to read the information displayed on the subject's computer screen, answered questions, handed out the second set of instructions "for the buyer," and gave them \$5.00 in cash.¹⁰ The instructions for the buyer included the following information: (1) Buyers are expected to use their \$5.00 in trading sessions, and profits or losses they experience during these trading sessions will be added to or subtracted from this fund (i.e., they will be paid the initial "capital" of \$5.00 plus profits or minus losses at the end of the experiment). (2) The trades will be repeated many times (subjects were not informed of the total number of trials). (3) On each trial (or "trading session"), a buyer is matched with two sellers and has to buy

¹⁰ Dollar figures in the experiments reported below should be translated to yen figures with the exchange rate of \$1.00 for ¥200 (see n. 2 above).

a commodity from one of those two sellers. (4) Each seller is also matched with two buyers. Thus, when one of the sellers completes a deal with another buyer, the subject has to trade with the remaining seller. (5) One of the sellers (seller C) will be available throughout the experiment; the other seller (A) will be replaced by another seller (F) in the middle of the experiment. (6) Each trial or trading session starts with sellers announcing the prices of the commodity they are selling. Then, comparing sellers' offering prices, buyers must decide from which seller to buy. (7) When the quality of the commodity is standard, the experimenter buys the commodity for \$1.40. The actual quality of the commodity sold by a particular seller may be above or below this standard. The quality of the commodity sold by the same seller may also vary from session to session. The seller is supposed to have some information about the quality of the commodity she is selling, but the information may not be very accurate. Thus, sometimes the commodity is overpriced or underpriced, and when that happens it is not clear if the pricing is intentional or not. The buyer resells the commodity she bought from a seller to the experimenter. The experimenter purchases the commodity from buyers at prices reflecting its true quality. Subjects make profits when their purchasing price is lower than the resale price (which accurately reflects the quality of the commodity); they suffer losses when the purchasing price is higher than the resale price. (8) After each trade, a roulette spins on the computer displays of the trading partners (since the sellers are computer responses, the roulette actually spins only on the buyer's display). When the roulette hits a "strike," the seller is given a chance to "extort" a predetermined amount of money from the buyer. When this happens, the seller decides whether or not to use this chance and "rob" the buyer. As trials proceed, the amount to be extorted increases.

After the instructions, subjects were given four practice sessions to get used to the computer display and the keyboard. The real experiment started after the practice sessions were over and all questions were answered. Subjects had two sellers to trade with (sellers A and C) during the first 20 sessions. The computer was programmed such that each of the two sellers would have the chance to extort money from the subject twice. The amount to be extorted from the buyer increased from \$1.20 (sessions 1-7), to \$1.90 (sessions 8-14), and to \$2.60 (sessions 15-20). When given the chance to extort, seller A always used it, robbing the above amounts from the subject each time. In contrast, seller C never used the chance to exploit. In this way we made seller C a trustworthy partner. The offering prices did not systematically differ between the two sellers. Both made offers at prices randomly determined between \$1.00 and \$1.40 (uniform distribution). The resale price of those commodities were independently and randomly determined within the range of \$1.30

and \$1.50. Thus, subjects made profits on most of their trades although they suffered from extortion by A.

The "untrustworthy" seller A was replaced by another seller, F, after the twentieth session. For the rest of the experiment (sessions 21–30), the subjects had a different set of sellers to choose from, seller C (the trustworthy seller) and a new seller F with whom subjects had no prior experience. The new seller, F, made offers within the range of \$1.00 and \$1.25, undercutting the offering price of C on most trials. The experiment lasted for 30 sessions altogether. After the thirtieth trial was over, subjects were asked to fill out a postexperimental questionnaire, were given debriefing information, and were paid the amount they had earned in the experiment.

Manipulation of Social Uncertainty

Social uncertainty was manipulated by keeping or removing the "extortion" opportunities during the last 10 sessions (after the untrustworthy seller A was replaced by a new seller, F). In the low uncertainty condition, the roulette was removed from the subject's display and the subject was told that the sellers would no longer be given the chance to extort. In this condition, neither C (the trustworthy seller) nor F (the new seller) could extort from the subject. In the high uncertainty condition, on the other hand, the roulette and the chance to extort was not only kept during the last 10 sessions but the money to be extorted was increased to \$4.00. However, the roulette never hit a "strike" during the last 10 sessions and neither A nor F was actually given a chance to extort from the subject. Thus, the two conditions were equivalent with respect to the actual behavior of A and C.

RESULTS

Manipulation Check

First, whether or not the uncertainty manipulation had the intended effect on subjects was examined by analyzing a postexperimental questionnaire item, "How strongly were you concerned with the possibility that your money might be taken away during roulette chances?" As expected, the main effect of uncertainty in a nationality \times uncertainty analysis of variance (ANOVA) was significant ($F[1,193] = 4.66$; $P < .05$) and subjects in the high-uncertainty condition were more concerned ($M = 3.68$ on a 5-point scale; $SD = 1.11$) than were those in the low-uncertainty condition ($M = 3.36$; $SD = 1.36$). Second, seller A was evaluated by the subjects as less desirable than seller C (2.50 vs. 3.39 on a five-point semantic differential scale in which desirable = 5 and undesirable = 1; $t[196] = 5.27$;

$P < .001$). These results suggest that our experimental manipulations succeeded in creating the intended effects.

Hypothesis 1

The untrustworthy seller A was replaced by a new seller, F, after the twentieth session. The new seller was programmed to offer lower prices than C's price. The lower price provides an incentive for the subject to trade with F unless she is committed to C, the trustworthy seller. The frequency of sessions on which the subject bought from C instead of from F during the last 10 sessions is thus treated as an indicator of commitment formation by the subject with C. As predicted in hypothesis 1, subjects bought from C more frequently in the high-uncertainty condition ($M = 6.70$ sessions; $SD = 2.67$) than in the low-uncertainty condition ($M = 4.77$ sessions; $SD = 2.36$), and the main effect of social uncertainty in the uncertainty (subject's nationality analysis of variance was highly significant [$F(1,196) = 29.19$; $P < .0001$]). An analysis of postexperimental questionnaire items further confirms this result. Subjects were asked, "To what extent were the following considerations important to you in deciding on transactions during the experiment?" Possible answers were (1) keeping the same seller as a trade partner in order to develop a trusting relationship with him or her or (2) trading with a partner who can be trusted not to take away your money even when he or she gets an opportunity to do so. An analysis of the average of the subjects' responses (on a five-point scale) revealed a significant main effect of uncertainty ($F[1,193] = 4.19$; $P < .05$). Subjects felt the need to develop a trusting relationship with C more strongly in the high-uncertainty condition ($M = 4.11$; $SD = 0.99$) than in the low-uncertainty condition ($M = 3.81$; $SD = 0.99$). Hypothesis 1 is thus clearly supported by the results of this experiment as in Kollock's (1994) study (see fig. 1).

Hypothesis 2

The results indicate no difference in the degree of commitment formation with C between the two groups of subjects (5.85 sessions among Japanese subjects vs. 5.62 sessions among the Americans; $F[1,196] = 0.41$, NS). Furthermore, the effect of uncertainty did not vary significantly between the two groups as indicated by the weak, nonsignificant interaction effect in the above ANOVA ($F[1,196] = 0.85$, NS). In addition, neither the main effect of nationality ($F[1,193] = 0.01$, NS) nor the uncertainty \times nationality interaction ($F[1,193] = 0.08$, NS) in the analysis of the postexperimental questionnaire items was significant. These results clearly support hypothesis 2.

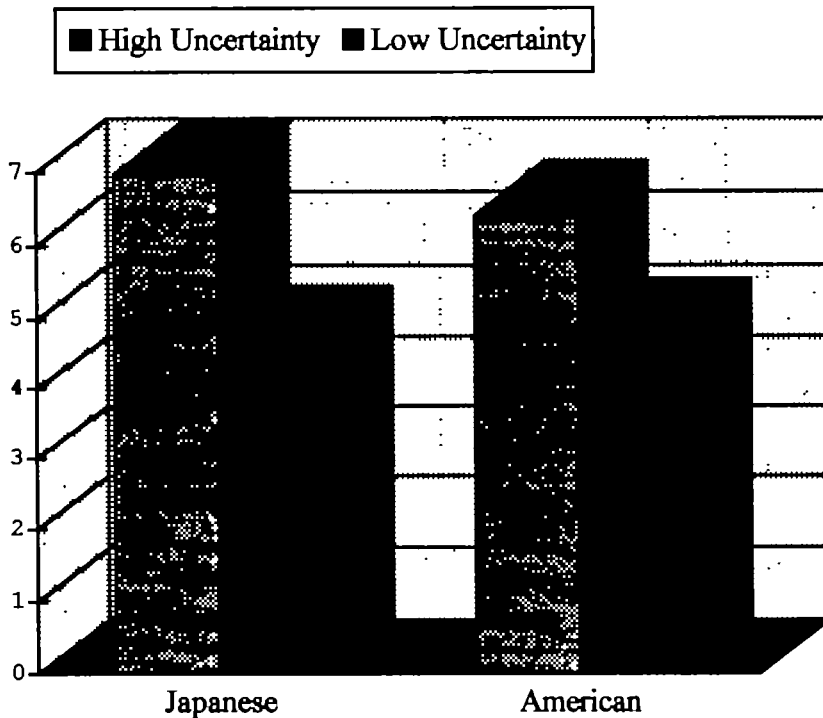


FIG. 1.—Commitment formation among Japanese and American subjects as a function of the level of social uncertainty.

EXPERIMENT 2

The results of the first experiment provide additional support for the proposition that social uncertainty facilitates commitment formation between particular exchange partners. The results further demonstrated that social uncertainty has comparable effects on commitment formation both among American subjects and Japanese subjects, at least in this experiment in which the two groups were matched on their levels of general trust. The purpose of the second experiment was (1) to replicate the effect of social uncertainty on commitment formation using a different experimental situation and to further enhance the generalizability of the proposition that social uncertainty facilitates commitment formation, (2) to test the other proposition critical to Yamagishi and Yamagishi's (1994) theory of trust, that low trusters, compared to high trusters, would be more strongly committed to a particular partner, and (3) to test the extent to which the above two propositions apply equally to American and Japanese subjects.

HYPOTHESIS 3.—Low trusters will form committed relations more strongly than high trusters.

HYPOTHESIS 4.—The effect of general trust on commitment formation predicted in hypothesis 3 is stronger when the level of social uncertainty is high than when it is low.

HYPOTHESIS 5.—The effect of social uncertainty and trust on commitment formation predicted above will be comparable for the two groups of subjects.

The second experiment was different from the first in the following respects. First, subjects' general trust was used as an experimental factor in this study. As mentioned earlier, American and Japanese subjects were matched with respect to their general trust levels in the first experiment, but trust level was not used as a factor in the data analysis due to the loss of the American subjects' individual trust scores in data shipment. Second, the experiment involved an abstract interdependent situation rather than the more concrete interdependence between sellers and buyers as used in the first experiment. Finally, two instead of one human subjects participated in each experimental group. Details concerning these procedural changes follow.

METHOD

Design and Subjects

The experiment involved a $2 \times 2 \times 2$ factorially crossed design (uncertainty = low vs. high; nationality of the subjects = American vs. Japanese; subject's level of general trust = low vs. high). All three factors are between-subjects factors. Subjects were recruited in a similar manner as in the first experiment. The subject pool in each country was divided into high trusters and low trusters at the median trust score (25 points) of the Japanese sample.¹¹ High trusters were those whose trust scores were 27 or higher; low trusters had trust scores of 23 or lower. From the Japanese subject pool, we drew 107 high trusters and the same number of low trusters; 100 American subjects from each trust level also participated in the experiment. However, a postexperimental analysis of the distribution of the subjects' trust scores revealed a systematic difference in the trust scores between the two countries; the American sample included a larger number of subjects whose trust scores were slightly below the cutting

¹¹ We used the median score of the Japanese sample for dividing the subject pool since the experiment was conducted in Japan before it was conducted in the United States and thus it was impossible to obtain the American median when the Japanese subject pool was divided.

score (i.e., below the Japanese median). This occurred since the population mean of the American pool was higher than the population mean of the Japanese pool (i.e., most Americans were high trusters according to the Japanese standard) and thus a majority of the low-trusting pool clustered slightly below the cutting score. We therefore decided to delete pairs in which at least one subject's trust score was only slightly below the cutting score (i.e., subjects with scores of 22 or 23) in order to make the American and Japanese subjects compatible on their trust scores.¹² This resulted in 186 Japanese (74 low trusters and 112 high trusters) and 141 American (42 low trusters and 99 high trusters) subjects.

General Procedures

The second experiment was conducted in 1994 in both Japan and the United States. As in the first experiment, subjects were led individually to a private room upon arrival and were handed a set of written instructions. They did not have the opportunity to see the other participants before, during, or after the experiment. They could sense the existence of other participants, however, due to the layout of the rooms. Although subjects were told that the experiment would be run in a group of several participants, each experimental group actually involved only two subjects. Two or more groups were run simultaneously to create the atmosphere that more than four subjects were in fact involved. When only one pair was run at a time, confederates were brought in to create the atmosphere of a larger group.

As in the first experiment, relevant information was displayed on the computer screen located in front of each subject, and the subjects entered their responses at the keyboard. Subjects' computers were controlled by the host computer located in the control room. The experiment consisted of 60 trials, but subjects were not informed of this. Subjects were asked to answer the postexperimental questionnaire, debriefed, and then paid for their participation. The whole experiment including instructions and the postexperimental questionnaire took about 80 minutes.

Subjects were each given \$5 (high-uncertainty condition) or \$3 (low-uncertainty condition) at the beginning of the experiment as "capital" to be used in the experiment.¹³ The take-home pay for the subject was this initial endowment plus the profit gained or minus the loss incurred in the

¹² The subjects' behaviors within a group (actually a pair) were strongly interdependent so that pairs rather than individuals were used in the analysis.

¹³ The amount of the initial endowment varied by the uncertainty condition, since the amount of profits to be earned in the experiment was expected to differ.

course of the experiment. Then, subjects performed five practice trials. The practice trials were programmed so that the subjects experienced all the possible cases described below. The experimental trials began after all the questions following the practice trials were answered. On each trial, the subject dealt with either a human participant or the computer. On the trial in which the subject dealt with the computer, she received an amount of money randomly determined within the range of 12¢ (or yen) and 27¢ (or yen) on the roulette displayed on the subject's computer screen. On the trial in which the subject interacted with another human participant, the two partners were each given 10¢. Then, one of the two was randomly selected for a chance to appropriate the partner's 10¢. If the one who was randomly selected used the option, she earned 20¢, and the partner's earnings on that trial were zero. If she did not use the option, each earned 10¢.

Whether the subject was to deal with the computer or with a human on the first trial was randomly determined with equal probability. (The subject was told that the assignment was random, but was not informed of the probability.) Since only two human subjects were included in each group, the assignments were synchronized. When they were assigned to interact with a human participant, they interacted with the same person. However, they were made to believe that at least four participants were involved and that the probability of interacting with the same person on the next interaction was small.

From the second trial on, whether the subject dealt with the computer or with a human participant was determined in the following manner. When the subject dealt with the computer on the previous trial, it was determined in the same manner as in the first trial. When the subject dealt with a human on the previous trial, the subject and the partner (actually the only two human subjects) were each given the choice of interacting with the same partner. If the two chose to interact with the partner from the previous trial, the two did in fact interact with each other on the new trial. If this happened, each was given 10¢ and one of the two was given a chance to appropriate the partner's money. When at least one of the two declined to interact with the same partner, then both faced the same situation as on the first trial. Each was randomly assigned to deal with the computer or to interact with a human participant. Since only two human subjects participated in each group, the same pair was formed every time the subject was assigned to interact with a human participant. However, subjects were made to believe that they might be matched with any one of several participants. Thus, the choice for the subject was interacting with the same partner or taking the chance of dealing with a computer (a sure gain of 12¢–27¢) or a "new partner."

Manipulation of Social Uncertainty

Social uncertainty in the current context is the expected loss caused by the self-serving behavior of the partner, which is a product of the probability of expected loss and the size of the expected loss. In the first experiment (as well as in Kollock's experiment), social uncertainty was manipulated by varying the probability of a loss being imposed. In the second experiment, however, social uncertainty was manipulated by varying the second component of the expected loss, that is, the size of the expected loss. This design change was introduced in order to test the generalizability of the previous findings concerning the effects of social uncertainty. We hoped to test the notion that the predicted effects of social uncertainty would occur whichever component of the expected loss (the probability estimate or the size of the loss) is used to manipulate social uncertainty. In the low-uncertainty condition, the loss caused by the partner's self-serving behavior (i.e., appropriation of the subject's money by the partner) is limited to 10¢. In the high-uncertainty condition, an extra loss of 50¢ was imposed in addition to the original 10¢ when the subject's partner used the appropriation option. The extra loss of 50¢ was explained to the subject as similar to the cost of a broken car window caused by the theft of a car stereo.

Opportunity Costs

The emancipation effect of trust presumed in the theory requires the existence of opportunity costs. The level of opportunity costs (i.e., the potential profits that could be obtained outside of the committed relationship, or the difference between the comparison level for the alternative (CL_{alt}) and the comparison level for the current relationship (CL ; see Thibaut and Kelley 1959) was set at a fairly high level in both uncertainty conditions. In a commitment relationship in which both partners refrain from using the appropriation chance, the subject earns 10¢ per trial. By choosing to interact with the same partner who has proven to be cooperative (not using the appropriation chance) in the past, the subject can expect a fairly sure 10¢. By leaving the security of the committed relationship, however, the subject is given a chance to deal with the computer, which gives a minimum of 12¢ and a maximum of 27¢. The difference between the sure 10¢ and the expected gain of 19.5¢ from the computer is the opportunity cost used in this experiment. In this situation involving an opportunity cost, high trusters who expect that most other participants will not use the appropriation chance even in a new encounter will be tempted to leave the committed relationship in search of a chance to meet the more lucrative partner (in this case, the computer). On the other hand, low trusters who expect that most other participants will use the appropriation chance

in a new encounter will hesitate to leave the already established security of a committed relationship in which the chance of getting duped is minimal.

RESULTS

Except for the results from questionnaire responses, all analyses use pairs as the statistical unit of analysis due to the interdependence of the behavioral responses of the subjects who participated in the study in dyads.

Manipulation Check

Subjects' responses to a postexperimental question—How strongly were you concerned with the possibility that your money would be taken away when the partner received a 'chance'?—were used to see if the manipulation of social uncertainty had the intended impact on the subjects. The main effect of social uncertainty in a nationality \times uncertainty \times trust ANOVA was highly significant ($F[1,320] = 10.14$; $P < .01$) proving the validity of the manipulation. Subjects in the high-uncertainty condition ($M = 3.37$; $SD = 1.13$ on a five-point scale) were more concerned with the possibility of getting duped by the partner than were those in the low-uncertainty condition ($M = 2.94$; $SD = 1.28$). The only other effect that was statistically significant was the main effect of trust ($F[1,320] = 17.48$; $P < .01$) indicating that low trusters ($M = 3.47$; $SD = 1.27$) were more concerned than were high trusters ($M = 2.96$; $SD = 1.16$). The main effect of trust is a testimony to the predictive validity of the trust scale used in the assignment of subjects.

Utilization of the Appropriation Chances

Before engaging in hypothesis testing, let us briefly present a descriptive overview of how subjects used the appropriation chances available to them. The overall relative frequency of the trials during which subjects who were given the appropriation chance actually did use it was .51 ($SD = .31$). That is, subjects in this experiment engaged in exploitative behavior about half of the time they were given the opportunity to do so. In a nationality uncertainty \times trust ANOVA, the main effect of uncertainty ($F[1,156] = 55.37$; $P < .0001$) was significant. Subjects' exploitative behavior was much more frequent in the low-uncertainty condition ($M = .67$; $SD = .27$) than in the high-uncertainty condition ($M = .33$; $SD = .30$). This effect of social uncertainty suggests that subjects who saw the consequence of their exploitative behavior as more serious to the partner (involving the extra loss of 50¢) refrained from engaging in such behavior. Although this might seem to weaken the manipulation of social uncer-

tainty, the result of the manipulation check presented above indicates that the manipulation of social uncertainty was strong enough despite the fact that subjects in the high-uncertainty condition refrained from engaging in exploitative behavior. No other effects were significant.

Hypothesis 1

The degree of commitment formation in this analysis is measured as the probability that an existing pair of traders on one trial is maintained on the next trial. (It is a measure of repeat interactions.) As predicted in hypothesis 1, subjects in the high-uncertainty condition ($M = .30$; $SD = .27$) formed committed relationships (in which they kept interacting with the same partner) more frequently than did those in the low-uncertainty condition ($M = .19$; $SD = .18$), and the difference was highly significant in a nationality \times uncertainty \times trust ANOVA ($F[1,156] = .83$; $P < .001$).¹⁴ Furthermore, the commitment intention index confirms the same pattern. In this analysis, a commitment intention index was calculated for each individual by dividing the number of trials in which a subject chose to interact with the same partner and did not act exploitatively by the number of trials in which the subject was given a chance to keep interacting with the same partner. The main effect of social uncertainty was again significant ($F[1,156] = 18.11$; $P < .0001$) subjects wanted to interact with the same partner who had not used the appropriation chance more strongly in the high-uncertainty condition ($M = .48$; $SD = .24$) than in the low-uncertainty condition ($M = .36$; $SD = .15$). Finally, subjects in the high uncertainty condition revealed a preference for maintaining committed relations on all of the five relevant questions:

1. How important was it during the experiment to continue trading with the same partner? ($F[1,320] = 11.01$; $P < .001$)
2. Did you think that dealing with the same partner was advantageous insofar as the partner did not use the 'chance'? ($F[1,320] = 21.99$; $P < .0001$)
3. How important was it for you to act in a way to improve the chance of being chosen by the same partner continuously? ($F[1,320] = 11.25$; $P < .001$)
4. How strongly did you try to trade with the same partner in order to build a trusting relationship? ($F[1,320] = 27.12$; $P < .0001$)

¹⁴ The distribution of the commitment index is highly skewed as is shown in the comparison of the mean and the standard deviation. Thus, the same analysis was conducted with a square root transformation. The results were similar to the original analysis ($F[1,156] = 6.13$; $P < .05$).

5. How strongly did you wish to trade with a person who would not utilize the 'chance'? ($F[1,320] = 33.$; $P < .0001$)

These results clearly support the hypothesis that social uncertainty facilitates commitment formation between particular exchange partners.

Hypothesis 2

As predicted, nationality of the subjects did not have a main effect ($F[1,156] = 2.44$; NS) on commitment formation ($F[1,156] = 2.00$; NS) or on commitment intention, replicating the results of the first experiment.

Hypothesis 3

As predicted in hypothesis 3, low trusters ($M = .27$; $SD = .24$) formed committed relationships more frequently than did high trusters ($M = .23$; $SD = .23$). The main effect of trust in a nationality \times uncertainty \times trust ANOVA analysis of the pairwise index of commitment formation was significant with the original data ($F[1,156] = 3.99$; $P < .05$) and with the square root transformation ($F[1,156] = 4.80$; $P < .05$). The same pattern was observed with the commitment intention index, but the main effect did not reach the significance level ($F[1,156] = .56$; NS).

Hypothesis 4

The predicted uncertainty \times trust interaction did not reach the significance level in the analysis of the pairwise index of commitment ($F[1,156] = 1.84$, NS, with the original data; $F[1,156] = 1.39$, NS, with the transformed data). The commitment intention index data reveal a marginal effect of the uncertainty \times trust interaction ($F[1,156] = 3.01$; $P < .09$). The pattern of the interaction shown in figure 2 is consistent with the hypothesis, both with the pairwise data and the individual-level data. Furthermore, analyses of simple main effects of trust in the two uncertainty conditions show that trust has a significant effect in the high-uncertainty condition ($F[1,156] = 4.76$; $P < .05$, with the pairwise original data; $F[1, 156] = 4.81$; $P < .05$, with the transformed data; $F[1,156] = 2.65$; $P < .11$, with the individual-level data), but not in the low-uncertainty condition ($F[1,156] = 0.25$, NS, with the pairwise original data; $F[1,156] = 0.63$, NS, with the transformed data; $F[1,156] = 0.57$, NS, with the individual-level data). These results provide additional support for hypothesis 4.

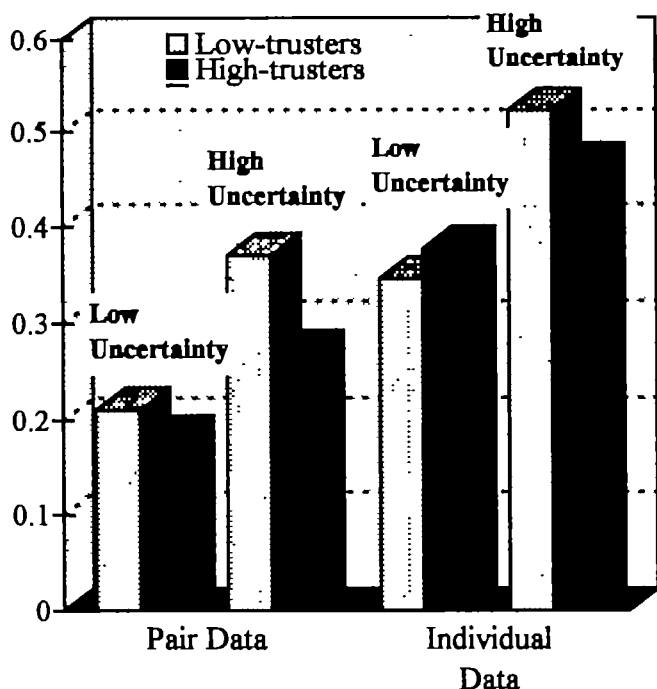


FIG. 2.—The degree of commitment formation and the strength of intentions to form a commitment relationship in experiment 2.

Hypothesis 5

As predicted, neither uncertainty nor trust significantly interacted with nationality in the nationality \times uncertainty \times trust ANOVA of commitment formation ($F[1,156]$ for the nationality \times uncertainty interaction = 1.75, NS; $F[1,156]$ for the nationality \times trust interaction = 0.06, NS) or for commitment intention ($F[1,156]$ for the nationality \times uncertainty interaction = 0.33, NS; $F[1,156]$ for the nationality \times trust interaction = 0.23, NS).¹⁵

CONCLUSION

The results of the experiments provide support for the two major propositions in the "emancipation" theory of trust. First, the proposition that social uncertainty promotes commitment formation was consistently sup-

¹⁵ Analysis with the transformed data also confirmed this conclusion.

ported in both experiments. Considering the substantial differences in the experimental design between our two experiments and between the current experiments and Kollock's (1994) experiment, it would be safe to conclude that support for this proposition is fairly robust. Commitment formation is a readily available and commonly used response among people facing socially uncertain situations, and the tendency to form stable relations with specific partners increases as the level of social uncertainty increases. In the first experiment, the level of social uncertainty was manipulated via a change in the likelihood of facing a risky situation in social interactions. In the second experiment, it was manipulated via a change in the potential loss caused by exploitative behavior of the interaction partner. The fact that the predicted effect emerged in both experiments provides a basis to argue for the validity of defining social uncertainty in this manner, that is, as the expected loss caused by an untrustworthy partner. Furthermore, social uncertainty had similar effects among American subjects and Japanese subjects in both experiments. This is consistent with the institutional view of culture discussed in the introduction. According to the widely shared view that characterizes Japanese culture as "collectivist," it is the "mind" of the Japanese people that makes them go along with the group. The results of these two experiments, however, demonstrate that once the relevant variables—degree of social uncertainty and the level of general trust—are controlled, American and Japanese subjects do not exhibit differences in their tendencies to voluntarily form committed relationships.

The second proposition that received experimental support is that the level of general trust is negatively related to the individual's tendency to form a stable relationship with a particular partner. Again, the predicted effect of general trust on commitment formation was not different among Americans and the Japanese. An additional proposition that the effect of trust would be stronger when the level of social uncertainty is high than when it is low also received experimental support, although the support was not as strong as it was for the other two propositions.

These results, taken together, provide further empirical support for the theory of trust proposed by Yamagishi and Yamagishi (1994). The theory emphasizes the role of general trust (trust in others in general) as an emancipator of people from the confines of safe but closed relationships. When a society as a whole is characterized by closed relations (as typically observed in collectivist societies), the one who is "emancipated" from a closed relation cannot find a better interaction partner since all the other relations are closed to "outsiders." The Japanese employment system among major companies for the past few decades is one of the best examples of such a "collectivist" society, although the situation is rapidly changing. Employment opportunities were almost completely closed to those making midlife

career changes, and thus opportunity costs for the employees of major companies were minimal. In such an environment developing a high level of general trust and becoming "emancipated" from the confines of established relations brings virtually no positive outcomes. Being highly trustful, expecting benign treatment from "strangers," in such an environment makes a person unrealistically optimistic. In contrast to this, we can think of a society in which better opportunities are abundant outside of the established relations. The American employment scene is closer to this ideal-type than to the previously presented collectivist ideal-type. Having a high level of general trust and not staying in the established relations despite better outside opportunities can have positive consequences in such an environment. Yamagishi and Yamagishi (1994) derived the hypothesis that Americans would have a higher level of general trust than the Japanese based on this reasoning and confirmed this hypothesis with survey data from a cross-societal questionnaire. The findings reported here provide support of a different kind—based on experimental methodology rather than survey research—and thus add to the validity of the theory.

The only anomaly observed in the data is the unexpected effect of the subject's nationality on the frequency of exploitative behavior. American and Japanese subjects were not different from each other in their tendencies to form committed relations, as expected. On the other hand, Japanese subjects showed a stronger tendency to behave exploitatively than did American subjects. Although this difference had not been predicted before the experiment, it is consistent with the institutional view of culture that underlies the current study. According to this view of Japanese society (Benedict 1946; Hechter and Kanazawa 1993; Yamagishi 1988a, 1988b), it is informal mutual monitoring and sanctioning rather than internalized moral values that insure that the Japanese will cooperate in achieving group goals. Yamagishi's (1988b) cross-societal experiment shows that once opportunities for monitoring and sanctioning are removed, the Japanese are in fact less cooperative in achieving group goals than are Americans. The experimental situation used in the current experiment resembles the one used in Yamagishi's (1988b) social dilemma experiment in this respect, and it is no wonder why the Japanese acted less cooperatively or more exploitatively than did the Americans.

The finding that nationality of the subjects did not play a major role in determining the level of commitment formation requires special attention. This finding suggests that cross-societal experimentation can be a powerful tool for exploring what have been relegated to the category of general "cultural differences." In this study, we started with a theoretical prediction that the level of commitment formation between particular partners is determined by the levels of social uncertainty and opportunity costs in

the environment and the level of general trust of the partners. The "cultural difference" that Japanese tend to form stable, long-term relations between particular partners or within particular groups should thus disappear once these three factors are experimentally controlled. And, this is exactly what happened in the two experiments reported above. If we still find a "cultural difference" even after experimentally controlling for the theoretically relevant variables, then we are further encouraged to look more deeply into the aspects of culture that are responsible for the "residual" differences. If we do not, then we do not need to revert to culture to explain the existing "cultural" differences. In brief, the goal of cross-societal experimentation is to demonstrate that "cultural" differences can be investigated more fully by experimentally controlling or manipulating the theoretically relevant variables. This is a different use of experimentation than is typical in cross-cultural experimentation, in which the primary goal is to demonstrate that cultural differences do exist; for example, to show that the Japanese act in a different manner than Americans do.

Finally, let us comment on the nature of the samples used in the experiment. It is clear that our samples do not represent any well-defined population within each society. This is a "problem" common to most experimental studies since it is practically impossible to bring a random sample drawn from a large population into the laboratory. What we have to keep in mind, however, is that experimental findings are not intended to be generalized beyond the laboratory in the descriptive sense. That is, descriptive generalization is not the purpose of theoretically informed experimental work. Instead, the goal of an such experimentation is theory testing. If a theory is disproved with a particular sample that is not representative of any specific population, the result at least tells us that there is an important set of hidden variables involved that interacts with the theoretical variables under study. This allows us to improve the theory. With experimental research, the critical kind of generalization is generalization of the theory or the specification of relevant variables that potentially interact with the primary theoretical variables, not generalization of the particular findings *per se*. For this purpose, replications with different experimental manipulations and situations are critical. We have demonstrated the generalizability of the propositions derived from Yamagishi and Yamagishi's (1994) theory of trust beyond one particular experimental setting.

REFERENCES

- Akerlof, G. 1970. "The Market for 'Lemons': Qualitative Uncertainty and the Market Mechanism." *Quarterly Journal of Economics* 84:488-500.
- Axelrod, R. 1984. *The Evolution of Cooperation*. New York: Basic.

- Benedict, R. 1946. *The Chrysanthemum and the Sword: Patterns of Japanese Culture*. Boston: Houghton-Mifflin.
- Cook, K. S., and R. M. Emerson. 1978. "Power, Equity and Commitment in Exchange Networks." *American Sociological Review* 43:721-39.
- Dawes, Robyn M. 1980. "Social Dilemmas." *Annual Review of Psychology* 31:169-93.
- Fukuyama, F. 1995. *Trust: The Social Virtues and the Creation of Prosperity*. New York: Free Press.
- Gambetta, Diego, ed. 1988. *Trust: Making and Breaking Cooperative Relations*. Oxford: Basil Blackwell.
- Gellner, E. 1988. "Trust, Cohesion, and the Social Order." Pp. 214-37 in *Trust: Making and Breaking Cooperative Relations*, edited by D. Gambetta. Oxford: Basil Blackwell.
- Hardin, R. 1990. "Trusting Persons, Trusting Institutions." Pp. 185-209 in *Strategy and Choice*, edited by R. J. Zeckhauser. Cambridge, Mass.: MIT Press.
- Hawthorn, G. 1988. "Three Ironies in Trust." Pp. 111-26 in *Trust: Making and Breaking Cooperative Relations*, edited by D. Gambetta. Oxford: Basil Blackwell.
- Hayashi, N. 1993. "From Tit-for-Tat to Out-for-Tat" (in Japanese with an English abstract). *Sociological Theory and Methods* 8:19-32.
- . 1995. "Emergence of Cooperation in One-Shot Prisoner's Dilemmas and the Role of Trust" (in Japanese with an English abstract). *Japanese Journal of Psychology* 66:184-90.
- Hayashi, N., N. Jin, and T. Yamagishi. 1993. "Prisoner's Dilemma Networks: A Computer-Simulation of Strategies" (in Japanese with an English abstract). *Research in Social Psychology* 8:33-43.
- Hayashi, C., T. Suzuki, G. Suzuki, and M. Murakami. 1982. *A Study of Japanese National Character* (in Japanese with an English summary), vol. 4. Tokyo: Idemitsu-shoten.
- Hechter, M., and S. Kanazawa. 1993. "Group Solidarity and Social Order in Japan." *Journal of Theoretical Politics* 5:455-93.
- Helper, S., and D. Levine. 1992. "Long-Term Supplier Relations and Product-Market Structure." *Journal of Law, Economics and Organization* 8:561-81.
- Jin, N., N. Hayashi, and H. Shinotsuka. 1993. "An Experimental Study of Prisoner's Dilemma Network: The Formation of Commitment in Selective Dyads" (in Japanese with an English abstract). *Japanese Journal of Experimental Social Psychology* 33:21-30.
- Kikuchi, M., Y. Watanabe, and T. Yamagishi. 1995. "The Ability of Accurately Predicting the Interaction Partner's Cooperativeness and the General Trust: An Experimental Study" (in Japanese). Paper presented at the Forty-third Annual Meeting of the Japanese Group Dynamics Association, September 23-24, Tokyo.
- Kiyonari, T., and T. Yamagishi. 1996. "Distrusting Outsiders as a Consequence of Commitment Formation" (in Japanese with an English abstract). *Japanese Journal of Experimental Social Psychology* 36:56-67.
- Kollock, P. 1994. "The Emergence of Exchange Structures: An Experimental Study of Uncertainty, Commitment, and Trust." *American Journal of Sociology* 100:313-45.
- Kosugi, M. 1996. "Trust and Gullibility: An Experimental Study." Paper presented at the Fiftieth Anniversary Conference of the Korean Psychological Association, June 27-29, Seoul, Korea.
- Kosugi, M., and T. Yamagishi. 1995. "Trust as a Cognitive Trait" (in Japanese). Paper presented at the Forty-third Annual Meeting of the Japanese Group Dynamics Association, September 23-24, Tokyo.
- Luhmann, N. 1979. *Trust and Power*. Chichester: Wiley.

- Messick, David M., and Marilynn B. Brewer. 1983. "Solving Social Dilemmas: A Review." Pp. 11-14 in *Reviews of Personality and Social Psychology*, vol. 4. Edited by Ladd Wheeler. Beverly Hills, Calif.: Sage.
- Schuessler, R. 1989. "Exit Threats and Cooperation under Anonymity." *Journal of Conflict Resolution* 33:728-49.
- Shapiro, D. L., B. H. Sheppard, and L. Cheraskin. 1992. "Business on a Handshake." *Negotiation Journal* 8:365-77.
- Thibaut, J. W., and H. H. Kelley. 1959. *The Social Psychology of Groups*. New York: Wiley.
- Vanberg, V., and R. Congleton. 1992. "Rationality, Morality, and Exit." *American Political Science Review* 86:418-31.
- Yamagishi, M., and T. Yamagishi. 1989. "Trust, Commitment, and the Development of Network Structures." Paper presented at the workshop, Beyond Bureaucracy, December 18-21, Hong Kong.
- Yamagishi, T. 1986. "The Provision of a Sanctioning System as a Public Good." *Journal of Personality and Social Psychology* 51:110-16.
- . 1988a. "Exit from the Group as an Individualistic Solution to the Public Good Problem in the United States and Japan." *Journal of Experimental Social Psychology* 24:530-42.
- . 1988b. "The Provision of a Sanctioning System in the United States and Japan." *Social Psychology Quarterly* 51:265-71.
- . 1988c. "Seriousness of Social Dilemmas and the Provision of a Sanctioning System." *Social Psychology Quarterly* 51:32-42.
- . 1992. "Group Size and the Provision of a Sanctioning System in a Social Dilemma." Pp. 267-87 in *A Social Psychological Approach to Social Dilemmas*, edited by W. B. G. Liebrand, D. M. Messick, and H. A. M. Wilke. Oxford: Pergamon.
- . 1995a. "Have Americans Really Become Distrustful?" Paper presented at the American Sociological Association Annual Meetings, August 19-23, Washington, D.C.
- . 1995b. "Social Dilemmas." Pp. 311-35 in *Sociological Perspectives on Social Psychology*, edited by K. S. Cook, G. Fine, and J. House. Boston: Allyn & Bacon.
- . 1995c. "Trust as Cognition, Trust as Reflection." Paper presented at the Second Workshop on Construction and Maintenance of Trust, September 9-10, Seattle.
- . 1995d. "Yakuzas, Lovers, Commitment, and Trust." Paper presented at the Workshop on Construction and Maintenance of Trust, February 25-26, New York.
- . 1996. "In-Group Bias and Culture of Collectivism." *Proceedings of the Korean Psychological Association Fiftieth Anniversary Conference*, pp. 119-33.
- . In press. "Trust and Establishment of Relationships" (in Japanese). In *Conflict Resolution*, edited by K. Ohbuchi. Tokyo: Nakanishiya Shuppan.
- Yamagishi, T., and N. Hayashi. 1996. "Selective Play: Social Embeddedness of Social Dilemmas." Pp. 363-84 in *Frontiers in Social Dilemma Research*, edited by W. B. G. Liebrand and D. M. Messick. Berlin: Springer-Verlag.
- Yamagishi, T., N. Hayashi, and N. Jin. 1994. "Prisoner's Dilemma Networks: Selection Strategy Versus Action Strategy." Pp. 233-50 in *Social Dilemmas and Cooperation*, edited by U. Schulz, W. Albers, and U. Mueller. Berlin: Springer-Verlag.
- Yamagishi, T., and H. Komiya. 1995. "Meanings and the Structure of Trust" (in Japanese with an English abstract). *INSS Journal* 2:1-59.
- Yamagishi, T., M. Watabe, N. Takahashi, N. Hayashi, and M. Yamagishi. 1996. "Trust and Commitment under Social Uncertainty" (in Japanese with an English abstract). *Research in Social Psychology* 11:206-16.
- Yamagishi, T., and M. Yamagishi. 1994. "Trust and Commitment in the United States and Japan." *Motivation and Emotion* 18: 9-66.

- . In press. "Trust and Commitment as Alternative Responses to Social Uncertainty." In *Networks in Japan*, edited by Mark Frunin. Cambridge: Cambridge University Press.
- Yamagishi, T., M. Yamagishi, N. Takahashi, N. Hayashi, and M. Watabe. 1995. "Trust and Commitment: An Experiment" (in Japanese with an English abstract). *Japanese Journal of Experimental Social Psychology* 35:23–34.
- Zucker, L. 1986. "Production of Trust. Institutional Sources of Economic Structure, 1840–1920." *Research in Organisational Behavior* 8:53–111.

Status Processes and Mental Ability Test Scores¹

Michael J. Lovaglia and Jeffrey W. Lucas
University of Iowa

Jeffrey A. Houser
Bowling Green State University

Shane R. Thye
University of South Carolina

Barry Markovsky
University of Iowa

The expected *consequences* of a score on an ability test can constrain individual performance. The authors predict that status processes, including status differences and the differences in rewards and costs that result, will produce differences in ability test scores between high-status and low-status individuals. In three controlled experiments, participants randomly assigned low status scored lower on a standard test of mental ability (the Raven Progressive Matrices) than did participants assigned high status. For both men and women, the difference in ability test score between low-status and high-status participants was about half a standard deviation. The results suggest the need to account for status differences in any attempt to measure mental ability accurately.

INTRODUCTION

Standardized tests of ability determine to a great extent who is admitted to elite institutions of higher education and increasingly who is hired and promoted in large organizations. Farkas et al. (1997) show that standardized test scores affect access to valued occupations and wages even when

¹ So many people contributed to this study that we can only thank a sample of them: Joseph Berger, Peter Burke, Thomas Fararo, Martha Foschi, Oscar Grusky, Kevin Leicht, John Skvoretz, Geoffrey Tootell, Henry Walker, Murray Webster, Jr., and David Willer. Direct correspondence to Michael Lovaglia, Department of Sociology, University of Iowa, Iowa City, Iowa 52242. E-mail: michael-lovaglia@uiowa.edu

controlling for other factors, such as education and work experience. The use of standardized test scores rationalizes the decision process, limiting the effect of personal bias on the part of decision makers. Unfortunately, reliance on standardized test scores has not equalized opportunities for members of disadvantaged social groups. African-Americans, for example, score lower than European-Americans on many tests of mental ability, including both college entrance examinations and IQ tests. Thus, use of standardized test scores for admissions and hiring decisions effectively excludes many disadvantaged group members. Farkas et al. (1997) conclude that to understand racial or ethnic inequalities in earnings, we must understand the social origins of group differences in standardized test scores.

We demonstrate how status processes that pervade society can lower the scores of disadvantaged group members on a standardized test of mental ability. We do so by extending well-developed theories of status processes to include performance on standardized tests. Then, we test our theory in a controlled, laboratory setting. Before describing the theory and its tests, we briefly review research on social structural conditions and social processes that affect standardized test scores.

Nature *and* Nurture

The nature of intelligence is a topic of flourishing research and debate. The terms "intelligence" or "mental ability" refer to an individual's capacity to understand complex ideas, to adapt to the environment, to learn, to reason, to solve problems, and to overcome obstacles by thinking about them (Neisser et al. 1996). A number of theorists have suggested new or expanded conceptions of mental ability (e.g., Damasio 1994; Gardner 1983; Sternberg 1985; Sternberg et al. 1995). However, the psychometric approach remains dominant. It uses standardized tests—the Stanford-Binet and the Wechsler IQ tests, for example—that presume to measure an underlying stable potential for high intellectual performance. For practical purposes such as school placement and personnel decisions, intelligence is a score on a standardized test (Scarr 1997).

Specifying the social factors that determine individual intelligence remains a critical unsolved problem in the scientific investigation of intelligence (Neisser et al. 1996). Ironically, research into the hereditary nature of intelligence provides ample evidence that social factors are important. The debate over whether heredity or environment is more important in the determination of intelligence has produced good empirical research. Nonetheless, the debate continues without hope of resolution.

Studies employing twins or adopted children have been used in attempts to disentangle the relative contribution to intellectual ability of

heredity and environment. A recent, well-designed, large-scale study comes down firmly in favor of heredity. The Minnesota study of twins reared apart (Bouchard et al. 1990) found that intelligence scores of monozygotic twins reared apart correlated about .70 while the scores of monozygotic twins reared together correlated about .80. That is, twins who lived apart in different environments were almost as similar in intelligence as twins who shared the same environment from birth. However, twin studies are not controlled experiments. A number of other factors could be responsible for the similarity between twins reared apart and twins reared together (Eysenk and Kamin 1980).

Adoption studies also support both nature and nurture. One startling conclusion drawn from these studies is that adopted children raised in the same family may be about as different from one another as children randomly selected from the population (Plomin and Daniels 1987). However, other studies compare the IQs of children living in deprived settings with the IQs of children adopted from deprived settings into more affluent homes. These studies generally report increased IQ for children placed in enriched settings and little evidence for IQ heritability (Schiff et al. 1978). In sum, evidence from twin and adoption studies supports the conclusion that both genetics and social factors play roles in determining individual intellectual ability.

For our purposes, it is sufficient to note the large role of environment in determining intellectual ability. Proponents of genetic determinism interpret the results of the Minnesota twin study (Bouchard et al. 1990) to mean that heredity is responsible for at most 70% of differences in intellectual ability. The environment, then, would account for at least 30%. The debate continues over the proper contribution of heredity and environment implied by these percentages. Other recent studies estimate a smaller role for heredity, a contribution of about 50% of the variation in IQ scores, suggesting a larger role for social factors (Chipuer, Rovine, and Plomin 1990; Devlin, Daniels, and Roeder 1997; Loehlin 1989; Rodgers, Rowe, and May 1994; Scarr and Weinberg 1978; Scarr, Weinberg, and Waldman 1993). If both heredity and environment make important contributions to individual intelligence, then the proportion that each contributes is not as important as identifying how those contributions are made.

The potential for cultural bias in standardized tests has been a major concern for several decades. Some standardized test items may be easier for privileged members of society to answer than for the less privileged. Whereas tests based on verbal and mathematical ability cannot be completely culture free, it is more difficult to make a case for cultural bias in nonverbal tests of abstract reasoning such as the Raven Progressive Matrices. Despite years of trying to eliminate cultural bias from standardized tests and increased education for African-Americans, they still score

lower than European-Americans on standardized tests—including IQ and scholastic aptitude tests. The difference remains substantial, around three-quarters of a standard deviation for IQ (10–12 IQ points) and two-thirds of a standard deviation for scholastic aptitude tests (Herrnstein and Murray 1994). The gap persists despite attempts to statistically control socioeconomic status and other social factors (Herrnstein and Murray 1994; Jensen 1992). If social factors are responsible for differences in test scores between social groups, then it is incumbent upon social scientists to identify those factors and demonstrate their impact on standardized test scores.

Moreover, differences in intellectual ability between groups may result from social factors even if individual differences in intelligence are largely inherited. The variation in test scores among individuals is in general much higher than the variation in scores between groups (Jensen 1980). For example, Neisser et al. (1996) found little evidence for genetic differences in intelligence between races. When social factors such as socioeconomic status are controlled, however, a substantial difference remains between ability test scores of African-Americans and European-Americans. Perhaps because specific social factors have not been demonstrated to produce substantial differences in ability test scores, the controversy continues over a genetic explanation for the difference between African-Americans and European-Americans.

Environmental Correlates of Ability Test Scores

Most of the research on social factors that could account for differences in ability test scores is correlational. Relatively little research has focused on social processes that could explain how social factors could produce test score differences. This section looks at social factors found to correlate with ability test scores. The following sections focus on social processes that could produce differences in test scores.

Some environmental factors have direct biological effects that are reasonably well understood. For example, poor nutrition during child development, environmental lead, and prenatal exposure to alcohol can all lower IQ scores (Pollitt et al. 1993, Needleman; Geiger and Frank 1985; Streissguth, Barr, and Sampson 1990). Factors such as nutrition, lead exposure, and prenatal alcohol exposure have direct biological effects on brain development that suggest an approach to improve social conditions: providing adequate nutrition or removing harmful agents from a child's environment eliminates the risk of low IQ from these causes.

Social factors that correlate with test scores often provide little insight into the process by which a social category influences individual test scores. For example, if racial differences in ability test scores are not due to

genetics (and there is no reason to believe they are), then why do African-Americans score lower than European-Americans? Socioeconomic status (SES) offers a partial explanation. African-Americans have disproportionately low SES, and low SES individuals do score lower on ability tests. The correlation between SES and IQ is about .4 (White 1982). However, test score differences between African-Americans and European-Americans remain when individuals of similar SES are compared (Loehlin, Lindzey, and Spuhler 1975).

Identifying social factors correlated with ability test scores also leaves open the question of causal mechanism. For example, one way that SES might alter test scores is that children from higher income families grow up in an enriched intellectual environment. Rodgers et al. (1994) found some evidence for the benefit of an enriched home environment in general and more specifically for the number of books owned by a child. However, the magnitude of the effect was neither large nor consistent across various ability test scores (Rodgers et al. 1994). Further, it is as easy to argue that high IQ leads to the acquisition of books as that acquiring books improves IQ.

It might help to look for the social processes that lead to differences in ability test scores rather than to differences in ability. Recall that, for practical purposes, mental ability is a score on a test. Ability is one determinant of an individual's score on an ability test. There are others. For example, Milofsky (1989) found that, in a suburban, predominately white, school district, psychologists spent twice as much time testing each student as did psychologists in an urban, predominately black, school district. Thus, even if individual psychologists treat African-American and European-American students identically, individual African-Americans have less time to complete the test and get less attention from testers. One way, then, that social factors such as race and SES can affect ability test scores is through the way tests are administered.

Expectancy Effects, Self-Esteem, and Self-Efficacy

Teachers' expectations affect their students' performance. Students try to fulfill their teachers' expectations and teachers' expectations bias evaluations of student performance. *Pygmalion* is the classic literary example of the process. Professor Henry Higgins's expectations for Eliza Doolittle transform her. A self-fulfilling prophecy operates whereby members of some groups are expected to be more competent than others, and the expectancy creates conditions that produce the expected result (Merton 1948).

In Rosenthal and Jacobson's ([1968] 1992) original study of the Pygmalion effect in the classroom, researchers led teachers to believe that some

of their students were likely to "bloom" intellectually during the coming school year. At the end of the school year, those students whom teachers expected to show greater intellectual improvement did show significantly greater gains on a mental ability test than did "nonbloomers." Surprisingly, researchers also found that teachers judged "nonbloomers" students unfavorably when they scored higher than expected. Rosenthal (1994) concluded that there are hazards to unexpected intellectual growth, an idea we will use later to develop our theory.

Despite the hundreds of studies that have reported expectancy effects in various social situations (Rosenthal and Rubin 1978), controversy over the importance of expectancy effects continues. The size of the effect remains unknown even among proponents (Rosenthal 1994). In addition, there is substantial opposition to the claim that teacher expectancies can influence learner intelligence. Snow (1995) reanalyzed the Pygmalion data and found only very small effects of expectancies on mental ability. He also points to the voluminous literature showing that mental abilities are not easily changed.

The expectations that individuals have for their own performance may also affect scores on mental ability tests. Social disadvantage could lead to lower ability test scores by adversely affecting self-esteem or self-efficacy. Self-esteem and especially self-efficacy are thought to improve performance by increasing the persistence with which individuals approach tasks and by reducing anxiety about possible failure (Bandura 1986). However, researchers have found no effect of self-esteem on achievement test scores (Maruyama, Rubin, and Kingsbury 1981) and only small self-efficacy effects on standardized test scores in a few studies (Multon, Brown, and Lent 1991).

The following sections develop and test a new theory that proposes a different social process to explain differences in ability test scores. Individuals in advantaged and disadvantaged groups hold different expectations about the personal *consequences* of an ability test score. Expectations about the personal consequences of a test score, rather than expectations about personal ability, may explain differences in ability test scores between social groups.

Status Processes and Rational Choice

We extend status characteristics theory to explain the difference in intelligence scores between advantaged and disadvantaged groups in society. Two elements of the status process work to the advantage of high-status individuals and the disadvantage of low-status individuals who take mental ability tests. First, previous social interaction as a high- or low-status individual may produce different expectations for performance on the test.

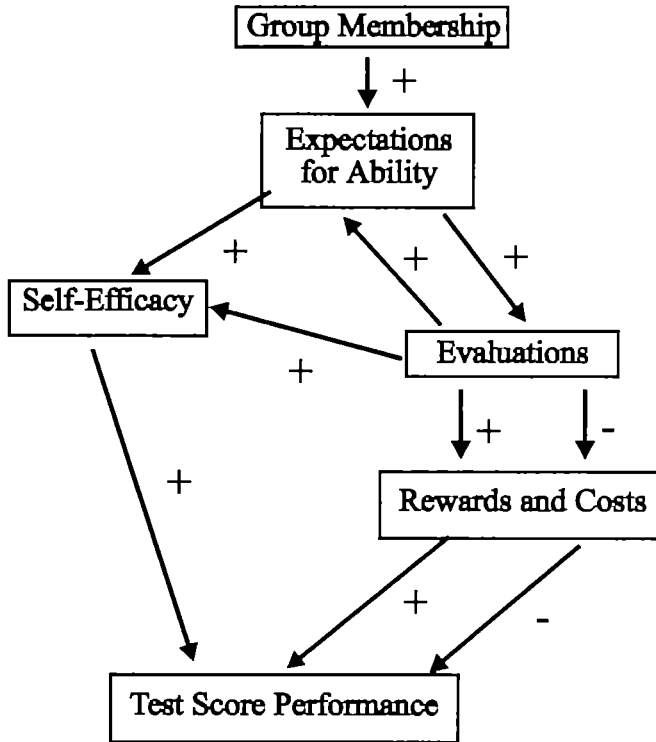


FIG. 1.—The status process

High-status individuals are evaluated more highly for their performances than are low-status individuals. Thus, in test situations, high-status individuals may have higher self-efficacy than low-status individuals. Second, according to status characteristics theory, status processes produce a social structure that provides rewards based on status. Those with high status come to expect high rewards for a competent performance. Those with low status expect not only low rewards but may anticipate punishment for competent performance that challenges the group's status hierarchy. Thus in some situations, it is in the interest of low-status individuals to underperform. Taking a standardized test may be one of those situations. A rational low-status individual may score lower on a standardized test rather than be penalized for a higher score.

The social process that produces status differences results in higher rewards and lower costs for identical performances depending on an individual's status (see fig. 1). Given different rewards and costs for identical performances, a rational choice perspective would suggest that increased

rewards and decreased costs would motivate a rational actor to do better when rewarded highly and worse when punished for a performance. We briefly introduce status characteristics theory, extend its scope to apply to individual performance on standardized tests, then show how the differential rewards and costs that result from the status process can further separate the IQ scores of high-status and low-status actors.

Status Characteristics Theory and Individual Performance

Status refers to an individual's standing in the hierarchy of a group based on the prestige, honor, and deference accorded her by other members. Status characteristics are features of individuals that influence group members' beliefs about each other. Different "states" of a status characteristic are assumed to have differential value, esteem, and honor. For example, in the United States, European-Americans are privileged over African-Americans. Race is a diffuse status characteristic because it carries with it expectations for competence in a wide variety of situations. Status characteristics can also be as specific as grade point average in high school or the score on a standardized test. Status characteristics help determine group members' relative status by altering expectations for competence that members hold for one another.

Status characteristics produce status rank through a chain of four logically connected assumptions (Webster and Foschi 1988):

1. A status characteristic becomes salient in a task situation if it differentiates among group members or is directly related to the task.
2. Salient status characteristics, even if not directly related to the task, will become relevant unless they are specifically dissociated from the task.
3. The effects of relevant status characteristics combine to form an aggregated performance expectation for each member.
4. Status rank is a direct function of the aggregated performance expectations of group members: The higher the aggregated performance expectation for a member, the higher is that member's status rank in the group.

The scope of status characteristics theory is confined to task-oriented groups where the contributions of all members are needed to accomplish some task. That is, status characteristics theory applies to groups where members are collectively oriented and task oriented. In groups meeting its scope conditions, the theory states that a status hierarchy will form

consistent with statuses that members possess in society at large.² High-status members (1) are given more opportunities to perform, (2) perform more often, (3) are given higher evaluations for their performances, and (4) have more influence over group decisions than do low-status members (Berger, Rosenholtz, and Zelditch 1980). Thus, status processes produce a self-fulfilling prophecy. Expectations for competence determine status rank, and high-status members are evaluated as more competent because they have high status. High evaluations lead in turn to higher rewards for high-status individuals (Berger, Fişek, et al. 1985).

Status characteristics theory also explains why low-status group members may be penalized for demonstrating they are more competent than their low status would suggest. Recall that teachers judged students unfavorably when the students violated teachers' expectations by performing at a higher level (Rosenthal 1994). Performance by high-status and low-status group members is perceived differently. While high-status individuals are given high evaluations for their performances, performances by low-status individuals are devalued and ignored despite their objective merit. Thus, while some low-status members may have very high ability, status hierarchies based on expectations for group members' ability are maintained as stable social structures. Very competent performances by low-status individuals do not produce comparable increases in expectations of their ability. Instead, unexpectedly competent performances by low-status individuals are seen as illegitimate (Ridgeway 1988; Ridgeway and Berger 1986, 1988) and subject to negative sanctions (Berger et al. 1998). Also, Ridgeway (1978) proposed and later demonstrated (Ridgeway 1981, 1982; replicated by Shackelford, Wood, and Worchel [1996]) that contributions of low-status group members are perceived to be selfishly motivated, while contributions of high-status group members are assumed to be group motivated. A high score on an ability test might well be negatively sanctioned if it were perceived as an illegitimate and selfish attempt to grab status. An expression used in the southeastern United States succinctly captures how group members feel about a competent performance by a low-status group member: Overachievement by those of low status is considered *uppity* and, therefore, subject to a number of social sanctions.

A story told by John Lamont, a successful African-American physicist, illustrates how displays of competence by low-status individuals can be sanctioned (Benjamin 1991). Lamont's father was a self-taught aeronauti-

² We present only those parts of status characteristics theory necessary to our argument. For more thorough exposition see Berger, Fişek, Norman, and Zelditch (1977), Markovsky, Smith, and Berger (1984), Webster and Foschi (1988), Berger, Fişek, and Norman (1989), or Berger, Norman, Balkwell, and Smith (1992).

cal engineer who invented a number of devices for use on airplanes during the 1930s and 1940s. However, given the prevailing racial climate, he supported his family working as a janitor at a gas company in Washington, D.C. When job opportunities opened up after the Depression, he decided to try for a better job. He told the foreman he would like to be a machinist. The foreman was skeptical of his ability to operate the machines. So Lamont's father demonstrated his ability as a skilled craftsman by operating the machines easily. His display of competence enraged the foreman who fired him on the spot from his job as a janitor.

Extending the Scope of Status Characteristics Theory

Status characteristics theory explains how group members expect superior performance from high-status members and evaluate their performance as superior even when performances by high- and low-status members are identical. However, the theory has not been used to predict the objective level of group members' performances. In particular, status effects on individual performance on standardized tests have been ruled out because such situations lack collective orientation and thus fall outside the scope of the theory. That is, because individual performance on standardized tests is independent of the contributions of other group members, the theory cannot predict that status information will alter the performance of test takers. To use the theory to predict differences in ability test scores, we must show how it can apply to individual performances.

Ridgeway and Walker (1995) note that status processes have been observed to constrain individual performances independent of actual ability. Hints of a role for status processes in performance on ability tests can be found in the research literature. Elizabeth Cohen and her colleagues have designed school programs to integrate students of diverse backgrounds in a cohesive classroom (Cohen 1986, 1993; Cohen, Lotan, and Leechor 1989). They succeed by carefully controlling status processes and by breaking down existing status distinctions (Cohen and Roper 1972, 1985; Rosenholtz 1985; Rosenholtz and Cohen 1985). An interesting by-product of the program is improved performance on standardized achievement tests for all students but especially for lower-status students (Cohen et al. 1989).

IQ gains made by children adopted into enriched environments have been found to fade by early adulthood (Scarr and Weinberg 1978). This has been seen as evidence of the genetic basis for intelligence (Herrnstein and Murray 1994). However, it also is possible that the IQ gains of young adoptees fade because status processes in school and work situations become more important as children age and counter the effects of an enriched home environment.

Steele and Aronson (1995) gave African-American and European-American students a test composed of items from the verbal portion of the Graduate Record Exam. In one condition, students were told the test measured their verbal ability. In another condition, students were told the test merely was a means of familiarizing them with verbal problems they might encounter. European-American students did equally well on the test in both conditions. In contrast, African-American students did worse when told the test measured their ability. This suggests that the status of African-Americans plays a role in their performance on standardized tests: their scores may drop when they know the results can be used to compare their performance with that of European-Americans. Our goal is to explain the mechanism behind such stereotype vulnerability.

We propose that status processes constrain individual performances when those performances are expected to have an impact on the relative status of the performer in the future (Lovaglia and Lucas 1997). Status makes a difference on an individual performance when the results of the performance have status value, that is, when the performance is expected to be used to determine status rank in future group interaction. According to status characteristics theory, status rank is a direct function of the aggregated expectations of group members for each other's competent performance on collective tasks. Those expectations are determined in part by individual performances. For example, achievement test scores produce general expectations of competence. We expect a person who scored 1600 on her combined math and verbal Scholastic Assessment Test (SAT) before entering college to be able to contribute more than a person who scored 750. Standardized test scores have a significant impact on the future academic and work careers of Americans. Thus, we propose that status processes will affect ability test scores in the United States. There are several ways this could occur.

Individuals' expectations for performance on a test could affect their performance directly. Here the mechanism suggested by status characteristics theory is similar to that proposed by self-efficacy research. Individuals who perceive themselves to be more capable of success on a test will persevere in trying to solve problems and experience less fear of failure. There is a major difference between status characteristics theory and self-efficacy with regard to expectations of competence. Self-efficacy has to do with *beliefs* or *conceptions* about personal capability. The implication is that perceptions of self-efficacy are consciously held. In contrast, status characteristics theory makes no assumption that expectations of competence are consciously held. Individuals may or may not be aware that they expect more competent performances from those with high status (Berger, Wagner, and Zelditch 1985). Still, the advantages accruing to high-status individuals on ability tests from increased perseverance and reduced anxi-

ety would probably produce only a small effect. We would not expect any greater effect for expectations of ability than the small effect of self-efficacy found by Multon et al. (1991).

William Foote Whyte ([1943] 1981) documented other ways that status processes affect individual performance in his classic *Street Corner Society*. Low-status gang members rarely beat gang leaders at bowling even when low-status members had superior bowling ability. If by chance a low-status member did beat one of the leaders, the low-status member could be taunted, ridiculed, and talked into losing a return match. This reaffirmed the status hierarchy of the group. An individual's bowling score is not dependent on a collective process and so would fall outside the scope of status characteristics theory, yet status processes seem to operate. It could be argued that there is a collective metatask—the task of maintaining the status hierarchy. A goal of bowling may be maintenance of status hierarchies, just as a goal of ability testing is the maintenance of status hierarchies. Thus, the same processes that affected bowling scores could affect scores on standardized tests. Low-status members underperformed at bowling because of the consequences of bowling well. In society, there also may be negative consequences for low-status individuals who do well on standardized tests.

Rational Choice

Individuals occupying different status ranks may come to expect quite different outcomes from the same performance on an objective, standardized test. These expectations may then affect how an individual performs on such a test. From a rational choice perspective, if people expect to receive large rewards for success on a test, they may do better on the test than they would if they expected a smaller reward. For example, some might expect a good score on the SAT test to lead eventually to a position as a prominent doctor or lawyer. But others, coming from different backgrounds, might expect a more modest reward, a steady job with the post office or as a teacher.

Members of different groups also may expect different costs to result from a score on an ability test. If people expect to pay substantial costs for success on a test, they may do worse than they would if they expected costs to be trivial. For some the costs of success might be trivial. A high score and going off to college entail little disruption in the life of the son or daughter of a doctor. Others may expect much higher costs. For example, a minority student who does well on a test and plans to go to college might be shunned by peers for trying to be “too white” (Fordham and Ogbu 1986; Steinberg, Dornbusch, and Brown 1992). African-Americans are particularly concerned about the costs of academic success. Arroyo

and Zigler (1995) showed that, for African-Americans, attitudes conducive to high academic achievement were associated with introjective depression and especially with concerns about losing the approval of others. Moreover, going away to college involves immersion in an alien culture, cut off from social support (Morris 1979; Blackwell 1981; Fleming 1981; Fordham 1988).

There also is evidence that teachers penalize low-status individuals for scoring higher on tests than teachers think they should. The original Pygmalion study found that intellectual ability is penalized when it violates expectations (Rosenthal and Jacobson 1968). If members of a disadvantaged group are expected to possess lower mental ability, then disadvantaged individuals who do well on ability tests face increased criticism. Rubovits and Maehr (1973) conducted a follow-up to the Pygmalion study that compared teachers' reactions to African-American and European-American students. African-American students thought to be "gifted" were criticized the most and given the least attention. European-American students thought to be "gifted" were praised the most and given the most attention. African-American students thought to be nongifted received almost as much praise and attention as did European-American students thought to be nongifted. Rosenthal's (1994) conclusion that a penalty is imposed on those who show unexpected intellectual ability holds true for race. African-Americans with high IQ scores were criticized more and praised less than other students, both black and white. That is, African-Americans were penalized for high scores on a standardized test.

We conclude from the above evidence that African-Americans not only expect to be penalized for a high score on standardized tests but actually do bear a cost for success. Fordham and Ogbu (1986) suggest that African-Americans grow up with a double message about intellectual achievement: (1) work twice as hard to get half as far, and (2) keep your head down, do not stand out. As a result, African-Americans experience ambivalence and dissonance toward intellectual effort and success (Fordham and Ogbu 1986). Thus, underperformance on an ability test represents an adaptive response by African-Americans. Getting a low score on a test would be a reasonable way to avoid those costs. However, it is possible, even likely, that low-status individuals would work as hard as anyone when taking an ability test. We propose that people taking an ability test try hard to get the best score possible without incurring an unacceptable cost. In the face of a possible severe penalty for success, low-status individuals should be extremely motivated to get just the right mediocre score.

Because the expected *consequences* of ability test scores have implications for status hierarchy formation, we propose an extension to status characteristics theory that allows its application to individual performances. If the extension proves valid, then an individual's rank in the

status hierarchy and the resulting rewards and costs associated with success on achievement tests may affect scores on such tests. We tested the following hypothesis for situations in which differences in an ability test score have implications for future work in a task group:

HYPOTHESIS. *If the performance expectations and reward expectations of ability test takers correspond to their status, then their performance is a positive function of their status.*

METHOD

We used an experimental approach to test our theory that the expected consequences of a score on an ability test partially determine the score that an individual receives. Researchers have recently pointed to the need for social research to control possible genetic or biological influences on social phenomena (Scarr 1997; Udry 1995). Eysenck (1995) and Turkheimer (1991) argue that the experimental investigation of differences in intelligence is now not only needed but possible. Experiments provide strong evidence for the causal direction of a relationship between two variables. We use laboratory experiments to investigate whether a fundamental social process—status hierarchy formation and maintenance—can produce differences in ability test scores. Experimental control allows us to pinpoint the cause of any test score differences we find and rule out competing genetic or biological explanations.

Along with its strengths, the experimental approach has disadvantages, as do all research designs. For maximum effectiveness, the experimental approach dictates, for example, that research participants are assigned randomly to the conditions thought to produce an effect. By randomly assigning participants to conditions, we obtain strong evidence that those conditions, and not something else, produced any observed differences in test scores. We can screen out extraneous systematic differences between experimental groups, while statistically controlling random differences. The disadvantage is that the kinds of conditions we can create in the laboratory are limited. For example, it is not possible to assign participants to a race randomly. It is not feasible to assign participants to any major social category associated with differences in ability test scores. We cannot randomly assign participants to different religions, to wealth or poverty, to different home environments. The problem of alternative plausible explanations that plagues other research methods is present in experiments where participants cannot be randomly assigned to conditions. For example, we could administer a standardized test to carefully matched African-American and European-American students. The results would likely show that African-American students scored lower, but we would not be any closer to finding out why. To effectively investigate the social process

that produces differences in ability test scores, we must assign participants randomly to either high-status or low-status conditions.

In a laboratory, we can create status differences that have social consequences. We can create conditions where high-status participants expect higher rewards and lower costs than low-status participants for a high score on an ability test. Once created, we can assign participants randomly to those conditions. Then we can administer a standardized test of mental ability and look for differences in test scores between conditions. That is, we can model the social process theorized to produce differences in ability test scores. If we find theoretically predicted differences, then we have strong evidence that the social process being modeled does produce differences in ability test scores.

The logic of experimental design is indirect. Results of experiments do not generalize to naturally occurring situations the way that survey questionnaire results do (Zelditch 1980). For example, surveys conducted before elections predict more or less accurately who will win the election. We know how accurate those surveys will be. The more representative of the population of voters is the sample used in the survey, the more accurate the results will be. Experiments do not work that way at all. It seems reasonable to ask, How can an experiment on white university undergraduates tell us anything about racial differences in test scores?

The logic of the experimental approach is commonly understood in other fields of research. Most of us are familiar with its application to medical research. For example, medical researchers cannot randomly assign people to be exposed to suspected cancer-causing agents, then wait to see who gets cancer. They use laboratory animals instead. Researchers paint tobacco tar on the skin of randomly selected mice, while randomly selecting other mice as controls. They then wait to see if mice painted with tobacco tar get cancer more often than do the control mice who were not exposed. When the painted mice get cancer much more often than the control mice, researchers are confident that the tobacco tar caused the cancer. The study is replicated under different conditions with different animals to investigate the process by which tobacco tar produces cancer. For example, dogs can be taught to smoke cigarettes and the cancer rates of smoking and nonsmoking dogs compared. However, as has often been said by executives of tobacco companies and their lawyers, studies on laboratory animals do not prove that smoking tobacco causes cancer in humans.

While laboratory studies have not proved that smoking tobacco causes cancer in humans, they have shed light on the physiological process that does. If the same physiological processes take place in humans and in particular laboratory animals, and if those processes have been shown to produce cancer in those animals, then we have reason to suspect that

cancer will result in humans as well. Added to the evidence of laboratory studies is the correlational evidence from human populations. People who smoke cigarettes get cancer at a much higher rate than nonsmokers do. Smokers who quit smoking for a number of years have a reduced chance of getting cancer. It is true that smoking tobacco has not been proved to cause cancer in humans. However, few people think that smoking is safe. The evidence linking smoking to cancer is overwhelming.

Laboratory studies of social phenomena employ the same logic. Experiments are best used to test theories of social processes (Mook 1983; Zelditch 1980). Experimental evidence is then used to develop better theories in a research program. In research programs, the relationship between theory and empirical investigation is reciprocal. Empirical research both tests theory and prompts theoretical development that then requires further tests. Research programs facilitate the cumulative growth of knowledge (Wagner and Berger 1985, 1986, 1993; Szmataka and Lovaglia 1996). We feel we understand a social process when the theory explaining it has been supported by many experimental tests of various aspects of the process. Then, when we understand an underlying social process, we gain confidence that theoretically derived predictions will hold in diverse situations that conform to the conditions specified by the theory (Webster and Kervin 1971).

We apply the logic of the experimental approach to the study of differences in ability test scores between groups. To do so we extended a well-tested theory of an underlying social process. Status characteristics theory explains how the process of status hierarchy formation operates to produce different ability test scores for high-status and low-status individuals. The theory has been supported by hundreds of tests of its various aspects.¹ We designed experiments to test the specific prediction that status processes can produce differences in ability test scores for high-status and low-status individuals. If the experiments find higher ability test scores for high-status than for low-status individuals, then we will gain confidence that we understand how status processes produce different ability test scores for advantaged and disadvantaged groups.

There is ample evidence that race operates as a status characteristic (Cohen and Roper 1972; Webster and Driskell 1978). Thus, if racial differ-

¹ See Berger, Wagner, and Zelditch (1985) for a review of the first 20 years of status characteristics and expectation states research. Berger et al. (1992) review more recent work and describe an extensive test. Cohen and Zhou (1991) found status characteristics to influence behavior in research and development teams that had existed for years in organizations. Recent theoretical developments relating status processes to other areas of social research have also been supported in experimental tests (Lovaglia and Houser 1996; Biernat and Kobrynowicz 1997; Troyer and Younts 1997; Willer, Lovaglia, and Markovsky 1997).

ences in society are characterized by the conditions specified by the theory to produce differences in ability test scores, then we will gain confidence that we understand how status processes produce racial differences in ability test scores. The theory specifies three general conditions for its application to racial differences in ability test scores: if (1) racial differences are such that African-Americans are expected to be less competent in a variety of work settings than are European-Americans, and (2) test scores have status consequences for future work in groups, and (3) African-Americans expect lower rewards and higher costs to result from success on tests of mental ability, then the theory proposes that status processes will produce test score differences between racial groups.

DESIGN

We created status differences in the laboratory and randomly assigned participants to either a high-status or low-status condition. Then we administered a standard test of mental ability to all participants. Differences in test scores between the high-status and low-status conditions represent strong evidence that status differences produced differences in test scores. We also collected data from participants for several relevant variables that can act as statistical controls.

The creation of status differences in the laboratory is the key to the experimental demonstration of the effect of status differences on ability test scores. Studies of naturally occurring status distinctions such as race inevitably confound social and hereditary causes. Steele and Aronson (1995) conducted experiments that show how test scores of African-Americans but not European-Americans are adversely affected by expected comparison with national norms. Thus, African-Americans are vulnerable to their social position, scoring lower when their scores could affect their social relations. Racial vulnerability appears to be a purely social effect, produced by test conditions in the laboratory. Randomly selected African-Americans in the comparison condition had lower scores than African-Americans in the no-comparison condition. However, even controlled laboratory experiments on race allow alternative explanations. The source of African-American vulnerability has yet to be determined. Would other students be as vulnerable in circumstances similar to those faced by African-Americans? Again we run into the random assignment problem. We cannot tell whether European-American students would respond similarly given similar experiences because we cannot randomly assign race.

Creating a Status Characteristic

To investigate whether status processes produce differences in ability test scores and to rule out alternative explanations, we created a status charac-

teristic in the laboratory as proposed by Ridgeway (1991; Ridgeway et al. 1995). We started with a status-neutral characteristic, handedness, that we felt would be an integral part of participants' identities. Because mental ability is a stable trait that has been reported to resist attempts to change it, we wanted to create a strong status effect. We felt that the status effect would be enhanced if participants identified with the characteristic, if the characteristic, like handedness, was part of them. We first established that handedness was a status-neutral characteristic. Survey responses from 384 undergraduates in the subject population showed no differences in expectations for the competence of left- and right-handed persons.⁴ We then set out to imbue a person's handedness with status value.

When students arrived for the study, they were asked if they were right- or left-handed, and a brightly colored wrist band was placed on the preferred wrist. (Handedness was later confirmed by a series of computerized questions.) Students were told they would be working in a group to solve difficult problems that required intense cooperation among group members. Computerized instructions informed students in one condition that research showed right-handedness to predict high ability in the kind of group work to be performed, while left-handedness predicted low ability.⁵ Right-handedness was further associated with several positive personal traits, while left-handedness was associated with several negative traits. The instructions explained that research has shown right-handers to be better in the kinds of group work to be done and that "certain psychological processes having to do with the left and right brain are thought to cause this effect." Students were told that research had shown left-handers to be more impulsive, disorganized, and prone to inattention. Further, students were told that research had shown some positive traits of left-handers, such as creativity, could increase the resentment of those who worked under them.⁶ In another condition, left handedness was associated with high ability and characterized positively, while right-handedness was associated with low ability and characterized negatively. Thus, students

⁴ The analyses are available on request from Michael Lovaglia.

⁵ The program to run the experimental setting is available from Michael Lovaglia on request.

⁶ We checked during debriefing to see whether the information on handedness was plausible. In only a few cases did students say they were suspicious to the point that they did not try their best on the ability test. Data for these students were removed before analysis. Much more common was the reaction by low-status individuals during debriefing who said they knew they were being discriminated against and tried harder on the test to prove discrimination would have no effect on them. These students generally scored below average on the test.

could be randomly assigned to a condition in which their handedness created expectations of either high or low status in the upcoming work with their group.

Altering Expectations of Rewards and Costs

We also wanted students to expect different rewards and costs to result from the status we assigned them. To do so, we set up three levels of occupational status, each with a different pay level, in their work groups: Supervisors were to be paid \$17 per hour; analysts, \$8; and menials, \$4.50. Students were informed of two criteria for assigning them to an occupation, both of which were purported to predict success at the task: (1) their status as a right- or left-hander and (2) their score on an aptitude test. High-status individuals who scored high on the aptitude test would be supervisors. High-status individuals who scored low on the test and low-status individuals who scored high on the test would be analysts. Low-status individuals who scored low on the test would be menials.⁷ Thus, while all students had a monetary incentive to do well on the test, the incentive was greater for high-status students.

On the cost side, students were told that low-status individuals were seldom appointed supervisor because of the conflict that sometimes erupted between low-status supervisors and other group members. They were told that low-status analysts were also harassed but not as severely. They were warned not to harass low-status group members. Further, students were warned against cheating and told that cheaters were often caught when low-status individuals scored abnormally high on the test. Thus, low-status students expected costs to result from a high score on the test while high-status students did not expect those costs.

In sum, students assigned high status expected higher rewards and lower costs to result from a high score on an aptitude test.⁸ We could then administer a standard ability test to determine whether students assigned high status would score higher than students assigned low status.

⁷ It can be argued that similar discrimination against African-Americans is effectively outlawed now in the United States. However, we propose that lower test scores result from the expectations of low-status individuals for lower rewards and higher costs. Thus the expectations of individuals, not the legality of discrimination, is the issue. We predict race, as a status characteristic, to produce differences in test scores to the extent that African-Americans expect lower rewards and higher costs to result from a high score.

⁸ We used several kinds of rewards and costs because we were unsure of our ability to produce a significant difference in ability test scores after a brief and relatively mild laboratory manipulation.

Raven Progressive Matrices Test

We chose to administer the Raven Progressive Matrices test to students for several reasons. First, it has been an accepted test of mental ability for many years (see Raven, Court, and Raven [1992] for a summary of standardization research as well as tests of reliability and validity). Second, Raven scores correlate highly with other measures of general mental ability and are considered closely related to Spearman's g , or general intelligence (Jensen 1992). Third, no reading is required, which makes the Raven less culture bound than some other tests of mental ability. Fourth, because it involves deciphering patterns, students' *a priori* expectations for their performance on verbal and quantitative problems would be less salient than on a scholastic aptitude test. Fifth, it is self-administered, and thus scores are not as subject to the bias of test givers as is the case for some intelligence tests.

Students were seated at computer terminals in individual lab rooms for the experiments. After approximately 15 minutes of computerized orientation about their status in the upcoming group task, students were given the Raven Progressive Matrices test. Upon completing the test, they were debriefed and paid.

Data Collection

During the computerized orientation, students answered several questions to confirm their identity as left- or right-handers. Data for those few who identified themselves as ambidextrous were removed from the study because assignment to the high- or low-status condition could not be determined. Students also reported their age, gender, high school grade point average (GPA), and score on the ACT test required for admission to the university. Other control variables of possible interest included father's and mother's levels of education and the estimated number of books in their home when they were 10 years old. During a post-test debriefing, students were questioned to determine whether they held expectations for their performance in group work that conformed to their assigned status.

STUDY 1

Altogether, 47 students (23 men and 24 women) took part in the first experiment. Participants were recruited from large survey courses at the University of Iowa. Data were not analyzed for three students because they reported that their suspicions regarding the status assignment caused them not to try their best on the test. That left data for 44 participants,

TABLE 1
ANALYSIS OF COVARIANCE, STUDY 1

Effects	<i>F</i>	<i>P</i>
Status	5.23	.028
Gender67	.417
ACT	5.97	.019
GPA02	.895
Status \times gender	3.29	.077

N = 44 men and women.

11 men and 11 women in each status condition. Except for two students of Asian descent, all were European-American.⁹

The mean Raven score for students in the high-status condition was 55.63 (*SD* = 3.03), significantly higher than the mean Raven score for low-status students (53.91; *SD* = 3.60; $t[42] = 1.72$; $P = .046$, one-tailed).¹⁰ A difference between conditions also appeared on one control variable. Low-status students reported a *higher* GPA in high school than did high-status subjects. Thus, overall means on the Raven test may understate the effect of status assignment.

Table 1 shows the results of analysis of covariance to estimate the effect of status assignment on the Raven score while controlling for students' ACT score, GPA, and gender. Note that the status effect remains ($F = 5.23$; $P = .028$). Score on a previous ability test also has an independent effect ($F = 5.97$; $P = .019$). Note also the marginal interaction between status and gender ($F = 3.29$; $P = .077$). We looked at the mean Raven scores of men and women separately to determine the nature of the interaction. Women, it appeared, were not affected by status assignment in the same way as were men. The mean Raven score for women was essentially the same in high-status ($M = 54.18$; *SD* = 3.49) and low-status ($M = 53.82$; *SD* = 2.52) conditions.

DISCUSSION

The discovery of gender differences concerned us because gender itself is a status characteristic (Pugh and Wahrman 1983; Johnson, Clay-Warner,

⁹ Although the lack of African-Americans avoids a possible confounding of race and status, we did not plan it. There were few African-American students in the subject pool and none volunteered for this study.

¹⁰ Results from a nonparametric test, the Mann-Whitney *U*, are similar ($Z = 1.92$; $P = .028$, one-tailed).

and Funk 1996). Most standard tests of intelligence are constructed to equalize scores of males and females. We did not expect any difference in scores between males and females on the Raven test. However, in study 1, high-status women scored about the same as did low-status men. The theoretical implications are interesting. If status has an effect on ability test scores, and gender is a status characteristic, then women should score lower than men on ability tests. If status differences do constrain ability test scores, then why do women not score lower than men on most mental ability tests?

To speculate about the effect of status processes on ability test scores for status groups in society, it is necessary to estimate how each group conforms to the conditions of the theory. Gender and race are status characteristics but much else as well. They are fundamental social categories in society. For example, we have shown that African-Americans conform well to the model. African-Americans are expected to be less competent than European-Americans on a variety of tasks and can expect lower rewards for the same level of ability. Perhaps most important, African-Americans face penalties, sometimes severe, for high scores on ability tests. Women are in a different situation. They may be expected to be less competent than men, and they can expect lower rewards. Women, however, do not face the penalties for high test scores that African-Americans do. Social conventions may require women to conceal their intelligence in certain situations, but a high-ability test score will not be penalized. The accepted path to success for women lies in doing well in school, getting high scores on tests, and going to an elite university to meet the right people. Women do not experience greatly increased costs relative to men until they enter the workforce full time, get married, and have children. Then the ambivalence toward personal achievement typical of African-Americans becomes apparent in women as well (Simon 1995). An example of the lower productivity of women compared to men can be seen in the lower publication rates for female scientists even though female scientists have IQs at least as high as male scientists (Cole 1987). Because women do not experience the same penalties for a high score on an ability test that African-Americans do, they are not likely to feel the profound ambivalence to a high score that the theory predicts for African-Americans.

Because Raven scores for women are not in general different from those of men we must examine the experimental situation for the cause of the emerging gender differences in study 1. We do so in study 3 when we investigate status differences in all-female work groups. First, however, we wanted to confirm that the differences found for men in study 1 were reliable.

TABLE 2
RAVEN SCORE REGRESSIONS, STUDY 2

Variable	b	SE b	P
Status	2.04	.99	.047
ACT32	.16	.059
GPA11	.40	.785

N = 40 men

STUDY 2

To avoid confounding gender with assigned status, we continued study 1 with only men to try to confirm the original result. Altogether, 43 men participated in the study. Data were discarded for three men, including two who identified themselves as ambidextrous and one who was suspicious of the status assignment and said he did not try hard on the ability test. Data for 40 men remain to be analyzed, 20 in each status condition. All men added to the study were European-American.

Male students assigned to the high-status condition scored significantly higher on the Raven test ($M = 56.20$; $SD = 2.88$) than did male students assigned low status ($M = 54.00$; $SD = 3.58$; $t[38] = 2.14$; $P = .020$, one-tailed).¹¹ No difference was found between status conditions on any control variable. We used multiple regression analysis to estimate the size of the effect of status assignment on Raven Score while controlling for students' high school GPA and ACT score. Table 2 shows that, controlling for these other measures of mental ability, the effect of status assignment on the Raven score remains about as large as the difference in overall means, that is, 2.04 points. Having established a significant difference in raw Raven scores between status groups, we transformed Raven scores into IQ scores ($M = 100$, $SD = 15$) for illustrative purposes only, to give a feel for the magnitude of the effect. Mean IQ for students assigned to

¹¹ Results from a nonparametric test, the Mann-Whitney U, are similar ($Z = 2.37$; $P = .009$, one-tailed). Data are censored at the high end of the scale for the Raven test. Because the test is designed for the general population, scores in our college student sample are quite high. The mean score for the high-status group is over 56. A perfect score is 60. Students who score a perfect 60 may have been able to score higher had the test allowed them. A possible result would be the underestimation of the difference between high- and low-status groups. However, the upper limit to Raven scores had little effect in this study. Only one subject in the high-status group and one in the low-status group scored a perfect 60 on the test.

the high-status condition was 120, while mean IQ of students assigned low status was 112, an eight-point difference in IQ (Raven 1990).

We conclude that randomly assigning male students to a high- or low-status position altered their scores on the Raven Progressive Matrices test. In study 1, however, no effect of the status assignment was found for female students.

STUDY 3

We altered the experimental setting to account for the status of women in society. The design of study 1 may have inadvertently imposed added costs on women in the high-status condition. Gender is a status characteristic. Women are expected to be less competent than men at a wide variety of tasks. But if women are low in status compared to men, the prospect of a leadership position in a mixed-sex work group would likely engender the same ambivalence in women as it would in individuals assigned low status. The gender composition of work groups was not specified in study 1. However, it is likely that many participants assumed work groups would be composed of both men and women. If so, Raven scores of female participants in study 1 might have been depressed in the high-status condition. Women may have expected higher costs to accompany a supervisory position than did men. That is, women in the high-status condition facing the prospect of being appointed supervisor may have expected the same kinds of harassment in the work group as low-status supervisors.¹²

If high-status women scored lower on the Raven test because their status as women would place them in an uncomfortable position as supervisor of a mixed-sex work group, then the solution is straightforward. For study 3, we changed the computerized instructions to make clear that women would work in groups composed only of women. To lessen male orientation in the group task, we changed the work group scenario from a business setting to a mental health treatment setting. The three levels of occupational status in work teams were changed: Psychotherapists were to be paid \$17.00 per hour; technicians, \$8.00; and orderlies, \$4.50.

We also added another cost to low-status students who scored high on the aptitude test. We informed students that low-status individuals with high scores on the test would be given the chance to become psychotherapists if they wrote and delivered a short speech to a professor and several graduate students. This increased the cost of a high test score for low-status students because many people fear public speaking. We also felt it would increase the involvement of low-status students, giving them a

¹² We thank Kevin Leicht for pointing this out.

TABLE 3
RAVEN SCORE REGRESSIONS, STUDY 3

Variable	b	SE b	P
Status	2.31	.92	.017
ACT31	.15	.052
GPA33	.40	.412

N = 40 women

chance to become supervisor. However, we hypothesized that despite any increased involvement, low-status students would still score lower on the mental ability test than would high-status students.

There were 44 women students in the study. Data for four students were discarded: one because she identified herself as ambidextrous, another had missing data, and two suspected the status assignment and said they did not try hard on the test. One woman, assigned to the high-status condition, said she was so upset at the unjust treatment of low-status students that she could not concentrate on the test. We included her results in the analysis, somewhat depressing the mean Raven score for high-status students. We analyzed data for 40 women, 20 in each status condition. Except for one Hispanic student assigned to the high-status condition, all were European-American.¹³

The results for women in study 3 replicate closely results for men in study 2. Women assigned high status scored significantly higher on the Raven test ($M = 54.95$; $SD = 2.93$) than did women assigned low status ($M = 52.35$; $SD = 3.39$; $t[38] = 2.59$; $P = .007$).¹⁴ We found no significant differences between women assigned high and low status on any of the control variables.

Table 3 gives results of the multiple regression analysis controlling for high school GPA and ACT score. Note that with these other indicators of mental ability controlled, the effect of status assignment remains about the same as the difference in the overall means, that is, 2.31 points. The size of the effect for women in study 3 was close to that for men in study 2, both in terms of raw Raven score and transformed IQ score. We conclude that the status assignment had similar effects on male and female students. Status processes produced a significant difference in ability test

¹³ As in study 1, although the lack of African-Americans avoids a possible confounding of race and status, we did not plan it (see n. 9 above).

¹⁴ Results from a nonparametric test, the Mann-Whitney U, are similar ($Z = 2.40$; $P = .008$, one-tailed).

score between participants assigned to high- and low-status conditions of about half a standard deviation in magnitude.

DISCUSSION

We developed a theory to explain why socially advantaged individuals score higher on ability tests than do disadvantaged individuals even though both take the test under identical conditions. We began with status characteristics theory and its propositions linking membership in socially advantaged or disadvantaged groups to inequalities in performance evaluations and rewards. In the theory, status characteristics that signal membership in advantaged or disadvantaged groups trigger expectations for ability that then produce observable differences in individual behavior. However, status characteristics theory had not been applied to individual performance on ability tests because the test situation seems to lack the collective social interaction necessary for status processes to occur.

Cohen's work using status processes to improve the academic performance of schoolchildren (Cohen et al. 1989) as well as Whyte's (1981) classic ethnographic account of status processes in a gang suggest that status processes do affect individual performances. We proposed an extension of the scope of status characteristics theory to include situations where individual performances have consequences for future interaction in work groups. These consequences of individual performances are what bring status processes to bear on those performances. For example, a high score on an ability test leads to higher expectations of ability. In status characteristics theory, status rank is a direct function of expectations of ability. Thus, high score on an ability test represents a bid for increased status. However, bids for increased status by low-status individuals may be seen as illegitimate and sanctioned. We proposed that if low-status individuals are penalized for high scores on an ability test, then it is in their interest to score lower. We predicted that the status-disadvantaged who expect low rewards and high costs to result from a high score on an ability test would score lower on the test than the status-advantaged.

Results of three experiments supported the hypothesis that participants randomly assigned low status would score lower on an ability test than would participants assigned high status. We used an accepted test of mental ability, Raven's Progressive Matrices. In study 1, men assigned low status had significantly lower Raven scores than did men assigned high status. However, no effect was found for women. In study 2, we increased the number of men who participated in the study to 40 to confirm that the effect found in study 1 was reliable. Study 2 found that men assigned high status scored about half a standard deviation higher on the Raven Progressive Matrices than did men assigned to low status. Thus, the mag-

nitude of the effect, while substantial, is below that found between African-Americans and European-Americans on such tests—about three-quarters of a standard deviation (Jensen 1980). However, the status effect found in study 2 resulted from about 15 minutes of instructions to participants that created mild differences in status and relatively small differences in rewards and costs.

Study 3 investigated why women responded differently than men to the conditions in study 1. We theorized that women may have been subjected to additional costs for a high score on a test. Women may have expected a high test score to result in leadership of a mixed-sex work group. Because of the status differences between men and women, women may have expected difficulties supervising men similar to the difficulties expected by participants assigned to the low-status condition. Thus, study 1 confounded gender with low status, explaining the result that high-status women scored about the same as participants assigned low status. In study 3, to resolve the problem, women expected to participate in work groups composed only of women. If the status of women were responsible for the lack of a result in study 1, then assigning women to work in same-sex work groups should solve the problem. In study 3, as predicted, women assigned to high status scored higher on the Raven Progressive Matrices than did women assigned to low status. The magnitude of the effect was comparable to that found for men in study 2, about half a standard deviation. While we did not anticipate the additional constraints on test scores for women produced by experimental conditions in study 1, resolution of the problem in study 3 provides independent confirmation of the effect of status processes on ability test scores.

Our studies raise several questions for future research. In our experiments, we looked for an effect of an entire status process. That is, a status process that includes both status differences and the differences in rewards and costs that result. Differences in status alone, independent of subsequent rewards and costs, may have a small effect on ability test scores. The effect would be similar to that of self-efficacy. There is some evidence that self-efficacy has a small effect on standardized test scores (Multon et al. 1991). It would be interesting to see if such an effect can reliably be produced in the laboratory by altering expectations that participants have for their ability. Also, penalties for a high score on an ability test may be the major factor in the low test scores we found for low-status individuals. Gender is a status characteristic, and women cannot expect rewards as high as men for a high score on an ability test, yet ability test scores for men and women are equal. Women, however, do not face any obvious additional penalties of a high score on an ability test. In contrast, African-Americans, who face a variety of additional costs for success, have substantially lower ability test scores. Thus, it may be the additional

penalties rather than lower expected rewards that produce most of the difference in test scores. Future experiments could systematically vary rewards and costs expected by participants to result from high-ability test scores.

Assessing the implications of our results requires care. We have discovered evidence that status processes—and the differential rewards and costs they generate—cause differences in ability test scores in the laboratory. Such evidence does not generalize directly to differences in test scores for status advantaged and disadvantaged groups in society. Rather, we have extended a well-developed and rigorously tested theory of status processes, then tested our extension. Because our experimental evidence supports the theory, we gain confidence that the theory can be applied wherever status-advantaged and -disadvantaged groups experience conditions required by the theory. To predict that members of a status-disadvantaged group will score lower on an ability test, the theory requires that (1) disadvantaged individuals are expected to have lower ability than advantaged individuals, (2) test scores have consequences for future work in groups, and (3) disadvantaged individuals expect a high test score to result in lower rewards and higher costs than do advantaged individuals.

Given the theoretical requirements for predicting which groups will score lower on ability tests, we can then use additional sources of evidence to determine how closely a particular group corresponds to the requirements of the theory. If the group faces conditions shown by the theory to produce differences in ability test scores, then we have reason to suspect that differences found for that group result from status processes. There is substantial evidence that African-Americans face conditions required by the theory to produce differences in ability test scores. First, African-Americans are expected to be less able on a variety of tasks requiring mental ability. That is, race is a status characteristic (Cohen and Roper 1972; Webster and Driskell 1978). The lower academic performance of African-Americans is widely known. Second, ability test scores determine who is admitted to elite institutions of higher education and who is given opportunities for high-status occupations. Thus, ability test scores have status consequences for future work in groups. Third, there is evidence that African-Americans can expect to be criticized for higher-than-expected test scores (Rubovits and Maehr 1973). In addition, African-Americans expect lower rewards and actual penalties to result from higher education. In a study of high school students, Mickelson (1990) found that, while African-Americans embraced the abstract value of education even more strongly than European-Americans, African-Americans were more pessimistic than European-Americans about the concrete rewards that would result. African-Americans can expect higher costs to result from academic achievement as well. Steele (1992) describes the high personal

and emotional costs involved when African-Americans attempt to complete a college degree, as well as the lack of expectations that a degree will offer any meaningful improvement in their lives. Because ability tests are strongly related to education, African-Americans may have similar expectations for ability tests. While more systematic evidence is needed of penalties imposed on African-Americans for high test scores, substantial evidence exists that African-Americans face the conditions required by the theory. If African-Americans do face theoretically required conditions, then we have reason to suspect that status processes lower their scores on ability tests.

CONCLUSION

We have shown how status processes, by altering expected rewards and costs, can affect scores on a standard test of mental ability. In three experiments, participants were given approximately 15 minutes of computerized instruction to create status differences and attendant expectations for rewards and costs. Participants in high-status and low-status conditions received exactly the same instructions. In all studies, the brief and relatively mild intervention produced differences in scores on a standard test of mental ability approximately half of a standard deviation in magnitude. Students randomly assigned to a high-status position scored higher on the Raven Progressive Matrices test than those assigned low status. The status effect remained after individual differences in mental ability were statistically controlled. Further research is needed to ascertain the effects of the more severe and long-term status processes that operate in society at large. Our results do suggest the necessity to account for status differences—and the expectations for rewards and costs that they produce—in any attempt to measure mental ability accurately.

REFERENCES

- Arroyo, Carmen, and Edward Zigler. 1995. "Racial Identity, Academic Achievement, and the Psychological Well-Being of Economically Disadvantaged Adolescents." *Journal of Personality and Social Psychology* 69:903–14.
- Bandura, Albert. 1986. *Foundations of Thought and Action*. Englewood Cliffs, N.J.: Prentice-Hall.
- Benjamin, Lots. 1991. *The Black Elite: Facing the Color Line in the Twilight of the Twentieth Century*. Chicago: Nelson-Hall.
- Berger, Joseph, M. Hamit Fişek, and Robert Z. Norman. 1989. "The Evolution of Status Expectations: A Theoretical Extension." Pp. 100–130 in *Sociological Theories in Progress: New Formulations*, edited by J. Berger, M. Zelditch, Jr., and B. Anderson. Newbury Park, Calif.: Sage.
- Berger, Joseph, M. Hamit Fişek, Robert Z. Norman, and David Wagner. 1985. "Formation of Reward Expectations in Status Situations." Pp. 215–61 in *Status Rewards*

- and *Influence*, edited by Joseph Berger and Morris Zelditch, Jr. San Francisco: Jossey-Bass.
- Berger, Joseph, M. Hamit Fişek, Robert Z. Norman, and Morris Zelditch, Jr. 1977. *Status Characteristics and Social Interaction: An Expectation States Approach*. New York: Elsevier.
- Berger, Joseph, Robert Z. Norman, James Balkwell, and Roy F. Smith. 1992. "Status Inconsistency in Task Situations: A Test of Four Status Processing Principles." *American Sociological Review* 57:843-55.
- Berger, Joseph, Cecilia Ridgeway, M. Hamit Fişek, and Robert Z. Norman. 1998. "The Legitimation and Delegitimation of Power and Prestige Orders." *American Sociological Review*, in press.
- Berger, Joseph, Susan J. Rosenholtz, and Morris Zelditch, Jr. 1980. "Status Organizing Processes." *Annual Review of Sociology* 6:470-508.
- Berger, Joseph, David G. Wagner, and Morris Zelditch, Jr. 1985. "Expectation States Theory: Review and Assessment." Pp. 1-72 in *Status, Rewards, and Influence: How Expectations Organize Behavior*, edited by Joseph Berger and Morris Zelditch, Jr. San Francisco, Calif.: Jossey-Bass.
- Biernat, Monica, and Diane Kobrynowicz. 1997. "Gender- and Race-Based Standards of Competence: Lower Minimum Standards but Higher Ability Standards for Devalued Groups." *Journal of Personality and Social Psychology* 72:544-57.
- Blackwell, James E. 1981. *Mainstreaming Outsiders: The Production of Black Professionals*. Bayside, N.Y.: General Hall.
- Bouchard, Thomas J., Jr., David T. Lykken, Matthew McGue, Nancy L. Segal, and Auke Tellegen. 1990. "Sources of Human Psychological Differences: The Minnesota Study of Twins Reared Apart." *Science* 250:223-28.
- Chipuer, Heather M., Michael J. Rovine, and Robert Plomin. 1990. "LISREL Modeling: Genetic and Environmental Influences on IQ Revisited." *Intelligence* 14:11-29.
- Cohen, Bernard P., and Xueguang Zhou. 1991. "Status Processes in Enduring Work Groups." *American Sociological Review* 56:179-88.
- Cohen, Elizabeth G. 1986. *Designing Groupwork: Strategies for the Heterogeneous Classroom*. New York: Teachers College Press.
- . 1993. "From Theory to Practice: The Development of an Applied Research Program." Pp. 385-415 in *Theoretical Research Programs: Studies in Theory Growth*, edited by Joseph Berger and Morris Zelditch, Jr. Stanford, Calif.: Stanford University Press.
- Cohen, Elizabeth G., Rachel A. Lotan, and Chaub Leechor. 1989. "Can Classrooms Learn?" *Sociology of Education* 62:75-94.
- Cohen, Elizabeth G., and Susan S. Roper. 1972. "Modification of Interracial Interaction Disability: An Application of Status Characteristics Theory." *American Sociological Review* 37:648-55.
- . 1985. "Modification of Interracial Interaction Disability." Pp. 350-78 in *Status, Rewards and Influence: How Expectations Organize Behavior*, edited by Joseph Berger and Morris Zelditch, Jr. San Francisco, Calif.: Jossey-Bass.
- Cole, Jonathan R. 1987. *Fair Science*. New York: Columbia University Press.
- Damasio, Antonio R. 1994. *Descartes' Error: Emotion, Reason, and the Human Brain*. New York: Putnam.
- Devlin, B., Michael Daniels, and Kathryn Roeder. 1997. "The Heritability of IQ." *Nature* 388:468-71.
- Eysenck, Hans J. 1995. "Can We Study Intelligence Using the Experimental Method?" *Intelligence* 20:217-28.
- Eysenck, H. J., and Leon Kamin. 1981. *The Intelligence Controversy*. New York: Wiley.
- Farkas, George, Paula England, Keven Vicknair, and Barbara Stanek Kilbourne.

1997. "Cognitive Skill, Skill Demand of Jobs, and Earnings among Young European American, African American, and Mexican American Workers." *Social Forces* 75: 913-37.
- Fleming, Jacqueline. 1981. "Stress and Satisfaction in College Years of Black Students." *Journal of Negro Education* 50:307-18.
- Fordham, Signithia. 1988. "Racelessness as a Factor in Black Students' School Success: Pragmatic Strategy or Pyrrhic Victory?" *Harvard Educational Review* 58:54-84.
- Fordham, Signithia, and John U. Ogbu. 1986. "Black Students' School Success: Coping with the Burden of 'Acting White.'" *Urban Review* 18:176-206.
- Gardner, Howard. 1983. *Frames of Mind: The Theory of Multiple Intelligences*. New York: Basic Books.
- Herrnstein, R. J. 1973. *IQ in the Meritocracy*. Boston: Atlantic Monthly Press.
- Herrnstein, R. J., and Charles Murray. 1994. *The Bell Curve: Intelligence and Class Structure in American Life*. New York: Free Press.
- Jensen, Arthur R. 1980. *Bias in Mental Testing*. New York: Free Press.
- . 1992. "Spearman's Hypothesis: Methodology and Evidence." *Multivariate Behavioral Research* 27:225-33.
- Johnson, Cathryn, Jody Clay-Warner, and Stephanie J. Funk. 1996. "Effects of Authority Structures and Gender on Interaction in Same-Sex Task Groups." *Social Psychology Quarterly* 59:221-36.
- Loehlin, John C. 1989. "Partitioning Environmental and Genetic Contributions to Behavioral Development." *American Psychologist* 44:1285-92.
- Loehlin, John C., Gardner Lindzey, and J. N. Spuhler. 1975. *Race Differences in Intelligence*. San Francisco, Calif.: Freeman.
- Lovaglia, Michael J., and Jeffrey A. Houser. 1996. "Emotional Reactions and Status in Groups." *American Sociological Review* 61:867-83.
- Lovaglia, Michael J., and Jeffrey W. Lucas. 1997. "Group Processes and Individual Scores on Standardized Tests: A Theoretical Note and Basis for Investigation." *Current Research in Social Psychology* 2:1-13. (available at <http://www.uiowa.edu/~grpproc>)
- Markovskiy, B., LeRoy F. Smith, and Joseph Berger. 1984. "Do Status Interventions Persist?" *American Sociological Review* 49:373-82.
- Mariyama, Geoffrey, Rosalyn A. Rubin, and G. Gage Kingsbury. 1981. "Self-Esteem and Educational Achievement: Independent Constructs with a Common Cause?" *Journal of Personality and Social Psychology* 40:962-75.
- Merton, Robert K. 1948. "The Self-Fulfilling Prophecy." *Antioch Review* 8:193-210.
- Mickelson, Roslyn Arlin. 1990. "The Attitude-Achievement Paradox among Black Adolescents." *Sociology of Education* 63:44-61.
- Milofsky, Carl. 1989. *Testers and Testing*. New Brunswick, N.J.: Rutgers University Press.
- Mook, Douglas G. 1983. "In Defense of External Invalidity." *American Psychologist* 38:379-87.
- Morris, Lorenzo. 1979. *Elusive Equality: The Status of Black Americans in Higher Education*. Washington, D.C.: Howard University Press.
- Multon, Karen D., Steven D. Brown, and Robert W. Lent. 1991. "Relation of Self-Efficacy Beliefs to Academic Outcomes: A Meta-Analytic Investigation." *Journal of Counseling Psychology* 38:30-38.
- Neisser, Ulric, Gwyneth Boodoo, Thomas J. Bouchard, Jr., A. Wade Boykin, Nathan Brody, Stephen J. Ceci, Diane F. Halpern, John C. Loehlin, Robert Perloff, Robert J. Sternberg, and Susana Urbina. 1996. "Intelligence: Knowns and Unknowns." *American Psychologist* 51:77-101.
- Needleman, H. L., S. K. Geiger, and R. Frank. 1985. "Lead and IQ Scores: A Reanalysis." *Science* 227:701-4.

- Plomin, Robert, and Denise Daniels. 1987. "Why Are Children in the Same Family So Different from One Another?" *Behavioral and Brain Sciences* 10:1-16.
- Pollitt, Ernesto, Kathleen S. Gorman, Patrice L. Engle, Reynaldo Martorell, and Juan Rivera. 1993. "Early Supplemental Feeding and Cognition." *Monographs of the Society for Research in Child Development* 58 (7), serial no. 235.
- Pugh, M. D., and Ralph Wahrman. 1983. "Neutralizing Sexism in Mixed-Sex Groups: Do Women Have to Be Better than Men?" *American Journal of Sociology* 88:746-62.
- Raven, John. 1990. *Raven Manual: Research Supplement 3*. Oxford: Oxford Psychologists Press.
- Raven, J. C., J. H. Court, and J. Raven. 1992. "Standard Progressive Matrices." *Raven Manual: Section 3*. Oxford: Oxford Psychologists Press.
- Ridgeway, Cecilia. 1978. "Conformity, Group-Oriented Motivation, and Status Attainment in Small Groups." *Social Psychology* 41:175-88.
- . 1981. "Nonconformity, Competence, and Influence in Groups: A Test of Two Theories." *American Sociological Review* 46:333-47.
- . 1982. "Status in Groups: The Importance of Motivation." *American Sociological Review* 47:76-88.
- . 1988. "Gender Differences in Task Groups: A Status and Legitimacy Account." Pp. 188-206 in *Status Generalization: New Theory and Research*, edited by Murray Webster, Jr., and Martha Foschi. Stanford, Calif.: Stanford University Press.
- . 1991. "The Social Construction of Status Value: Gender and Other Nominal Characteristics." *Social Forces* 70:367-86.
- Ridgeway, Cecilia, and Joseph Berger. 1986. "Expectation, Legitimation, and Dominance Behavior in Task Groups." *American Sociological Review* 51:603-17.
- . 1988. "The Legitimation of Power and Prestige Orders in Task Groups." Pp. 207-31 in *Status Generalization: New Theory and Research*, edited by Murray Webster, Jr., and Martha Foschi. Stanford, Calif.: Stanford University Press.
- Ridgeway, Cecilia L., Elizabeth Boyle, Kathy Jean Kuipers, and Dawn Robinson. 1995. "Interaction and the Construction of Status Beliefs." Paper presented at the annual meeting of the American Sociological Association, August, Washington, D.C.
- Ridgeway, Cecilia L., and Henry A. Walker. 1995. "Status Structures." Pp. 281-310 in *Sociological Perspectives on Social Psychology*, edited by Karen S. Cook, Gary Alan Fine, and James S. House. Boston: Allyn & Bacon.
- Rodgers, Joseph L., David C. Rowe, and Kim May. 1994. "DF Analysis of NLSY IQ/Achievement Data: Nonshared Environmental Influences." *Intelligence* 19:157-77.
- Rosenholtz, Susan J. 1985. "Modifying Status Expectations in the Traditional Classroom." Pp. 445-70 in *Status, Rewards and Influence: How Expectations Organize Behavior*, edited by Joseph Berger and Morris Zelditch, Jr. San Francisco, Calif.: Jossey-Bass.
- Rosenholtz, Susan J., and Elizabeth G. Cohen. 1985. "Activating Ethnic Status." Pp. 430-44 in *Status, Rewards and Influence: How Expectations Organize Behavior*, edited by Joseph Berger and Morris Zelditch, Jr. San Francisco, Calif.: Jossey-Bass.
- Rosenthal, Robert. 1994. "Interpersonal Expectancy Effects: A 30-Year Perspective." *Current Directions in Psychological Science* 3:176-79.
- Rosenthal, Robert, and L. Jacobson. (1968) 1992. *Pygmalion in the Classroom*. New York: Irvington.
- Rosenthal, Robert, and D. B. Rubin. 1978. "Interpersonal Expectancy Effects: The First 345 Studies." *Behavioral and Brain Sciences* 3:377-86.

- Rubovits, Pamela, and Martin L. Maehr. 1973. "Pygmalion in Black and White." *Journal of Personality and Social Psychology* 25:210-18.
- Scarr, Sandra. 1997. "Behavior-Genetic and Socialization Theories of Intelligence: Truce and Reconciliation." Pp. 3-41 in *Intelligence, Heredity, and Environment*, edited by Robert J. Sternberg and Elena Grigorenko. New York: Cambridge University Press.
- Scarr, Sandra, and Richard A. Weinberg. 1978. "The Influence of 'Family Background' on Intellectual Attainment." *American Sociological Review* 43:674-92.
- Scarr, Sandra, Richard A. Weinberg, and Irwin D. Waldman. 1993. "IQ Correlations in Transracial Adoptive Families." *Intelligence* 17:541-55.
- Schiff, M., M. Duyme, A. Dumaret, J. Stewart, S. Tomkiewicz, and J. Feingold. 1978. "Intellectual Status of Working-Class Children Adopted Early into Upper-Middle-Class Families." *Science* 200:1503-4.
- Shackelford, Susan, Wendy Wood, and Stephen Worchel. 1996. "Behavioral Styles and the Influence of Women in Mixed-Sex Groups." *Social Psychology Quarterly* 59:284-93.
- Simon, Robin. 1995. "Gender, Multiple Roles, Role Meaning, and Mental Health." *Journal of Health and Social Behavior* 36:182-94.
- Snow, Richard E. 1995. "Pygmalion and Intelligence?" *Current Directions in Psychological Science* 4:169-71.
- Steele, Claude M. 1992. "Race and the Schooling of Black Americans." *Atlantic Monthly* 69 (April).
- Steele, Claude M., and Joshua Aronson. 1995. "Stereotype Threat and the Intellectual Test Performance of African-Americans." *Journal of Personality and Social Psychology* 69:797-811.
- Steinberg, Laurence, Sanford M. Dornbusch, and B. Bradford Brown. 1992. "Ethnic Differences in Adolescent Achievement." *American Psychologist* 47:723-29.
- Sternberg, Robert J. 1985. *Beyond IQ: A Triarchic Theory of Human Intelligence*. New York: Cambridge University Press.
- Sternberg, Robert J., R. K. Wagner, W. M. Williams, and J. A. Horvath. 1995. "Testing Common Sense." *American Psychologist* 50:912-27.
- Streissguth, Ann P., Helen M. Barr, and Paul D. Sampson. 1990. "Moderate Prenatal Alcohol Exposure: Effects on Child IQ and Learning Problems at Age 7½ Years." *Alcoholism: Clinical and Experimental Research* 14:662-69.
- Szatmari, Jacek, and Michael J. Lovaglia. 1996. "The Significance of Method." *Sociological Perspectives* 39:393-415.
- Troyer, Lisa, and C. Wesley Younts. 1997. "Whose Expectations Matter? The Relative Power of First- and Second-Order Expectations in Determining Social Influence." *American Journal of Sociology* 103:692-732.
- Turkheimer, Eric. 1991. "Individual and Group Differences in Adoption Studies of IQ." *Psychological Bulletin* 110:392-405.
- Udry, J. Richard. 1995. "Sociology and Biology: What Biology Do Sociologists Need to Know?" *Social Forces* 73:1267-78.
- Wagner, David, and Joseph Berger. 1985. "Do Sociological Theories Grow?" *American Journal of Sociology* 90:697-728.
- . 1986. "Programs, Theory and Metatheory." *American Journal of Sociology* 92:168-82.
- . 1993. "Status Characteristics Theory: The Growth of a Program." Pp. 23-63 in *Theoretical Research Programs: Studies in Theory Growth*, edited by Joseph Berger and Morris Zelditch, Jr. Stanford, Calif.: Stanford University Press.
- Webster, Murray, Jr., and James E. Driskell, Jr. 1978. "Status Generalization: A Review and Some New Data." *American Sociological Review* 43:220-36.
- Webster, Murray, Jr., and Martha Foschi. 1988. *Status Generalization*. Stanford, Calif.: Stanford University Press.

- Webster, Murray, Jr., and John B. Kervin. 1971. "Artificiality in Experimental Sociology." *Canadian Review of Sociology and Anthropology* 8:263-72.
- White, Karl R. 1982. "The Relation between Socioeconomic Status and Academic Achievement." *Psychological Bulletin* 91:461-81.
- Whyte, William Foote. (1943) 1981. *Street Corner Society: The Social Structure of an Italian Slum*. Chicago: University of Chicago Press.
- Willer, David, Michael J. Lovaglia, and Barry Markovsky. 1997. "Power and Influence: A Theoretical Bridge." *Social Forces* 76:571-603.
- Zelditch, Morris, Jr. 1980. "Can You Really Study an Army in a Laboratory?" Pp. 531-39 in *A Sociological Reader on Complex Organizations*, 3d ed. Edited by Amitai Etzioni and Edward W. Lehman. New York: Holt, Rinehart & Winston.

Book Reviews

Coffee and Power: Revolution and the Rise of Democracy in Central America. By Jeffery M. Paige. Cambridge, Mass.: Harvard University Press, 1997. Pp. xv+432. \$45.00.

William Roseberry
New School for Social Research

In this impressive and important contribution to historical and political sociology, Jeffery M. Paige examines the 1980s political upheavals in Central America in terms of the distinct ways in which coffee elites in particular countries responded to social movements that threatened their positions and powers, both in the 1980s and in earlier moments of crisis. Concentrating on El Salvador, Nicaragua, and Costa Rica, Paige sketches the history of coffee cultivation in each country and explores the different class structures through which coffee was grown, processed, and commercialized. He then discusses two political conjunctures in each country—the country-specific crises in the first half of this century and the regionwide mobilizations of the 1980s—and argues that the specific resolution of the earlier crisis structured, socially and ideologically, the politics of the later conjuncture. In El Salvador, for example, the economic depression of the 1930s provided a seedbed for political opposition, including a communist movement. The movement's planned 1932 revolt served as a pretext for a massacre by government forces that consolidated the power of a coffee elite highly dependent upon military rulers who, together, viewed social movements and reform initiatives as a threat. The U.S. occupation of Nicaragua in the early 20th century foreclosed the liberal revolution that brought coffee elites to power in other Central American states, fostered the nationalist (not socialist) opposition of Sandino, and nurtured the familistic rule of the Somozas. In Costa Rica, a social reformist movement in the 1930s and 1940s, including the participation of communist activists, provoked a crisis and a brief revolution, the result of which was the installation of a social democratic regime at midcentury. Small producers and workers were incorporated within a welfare state, and democratic procedures and guarantees (including the abolition of a standing army) provided means for social movements to press grievances.

In outlining the class structure in each country, Paige is especially interested in the roles of the coffee sectors and their elites. He pursues two dimensions: first, the organization of production and characteristic relations between nonproducers and producers, and, second, a distinction between two segments of the coffee elite—the agrarians and the agroindus-

Permission to reprint a book review printed in this section may be obtained only from the author.

trial elite. The agrarians include the growers themselves; their power rests in land ownership and control of the actual processes of coffee cultivation. The agroindustrials control processing and finance and are more likely to dominate the export trade. While individuals and families may fit in both segments, they will owe their primary source of wealth to one. More important, the distinction marks two different productive processes ("agricultural" and "industrial") associated with coffee, leading to distinct interests and political stakes. Paige's understanding of a segmented coffee elite informs the class analysis itself, but Paige is most interested in exploring the changing balance between these two elite sectors, suggesting that the agrarians' dependence on repressive labor regimes leads them to support authoritarian political forms and outcomes while the agroindustrials are more likely to support and promote democratic regimes. The agroindustrials, in short, are bourgeois, and the agrarians cannot be. To the extent that the agroindustrials remain allied with the agrarians, as in El Salvador, repressive and authoritarian politics are pursued. Only where the agroindustrials firmly break with the agrarians, as in Costa Rica in 1948, is a democratic opening possible.

One will recognize here the influence of Barrington Moore, and Paige suggests the limited applicability of the "Moore thesis" to Central America—limited because even when the agroindustrial bourgeois embrace democratic forms they cannot see beyond the narrow confines of their sectoral interests and hesitate to open those forms to include the working classes. Paige contends that the agroindustrials have only broken with the agrarians and moved toward democratic politics under pressure from socialist movements from below.

An important feature of the book is that Paige supplements the structural analysis of political processes with extensive interviews with members of the coffee elite in each country. In his analysis of what he calls their "narratives," he explores the historical and political visions of the elite—their views of the coffee sector and its place within their country's history, of themselves as an elite and a political force, of past political conjunctures, and of the crisis of the 1980s. He is especially good at exploring differing elite visions, linked (especially in El Salvador and Nicaragua) to his agrarian/agroindustrial distinction. And he explores the structured silences in elite discourses with skill. In his presentation and interpretation of narratives, Paige retains full authorial control, summarizing positions and stories with occasional short quotes from individuals. We never read a full story or claim from one of the interviewees that is then subjected to interpretation. This is regrettable: the interviews were conducted during political ferment and crisis, and they provide important insight into the perceptions and ideology of central economic and political actors. One hopes Paige explores these materials more fully and substantively in another setting.

Paige's book is a fundamental contribution to political sociology, blending theoretical sophistication with deep historical knowledge of the specific cases considered. It is also a necessary text for an understanding

of the recent history of Central America. Its structural analysis is sharply focused and illuminating; its examination of elite narrative and ideology lends the book a subtlety often missing from structural analyses alone.

Restructuring Networks in Post-Socialism: Legacies, Linkages, and Localities. Edited by Gernot Grabher and David Stark. Oxford: Oxford University Press, 1997. Pp. x+349. \$75.00.

Ákos Róna-Tas
University of California, San Diego

This volume is a milestone in the market transition literature. Not all the pieces are equally brilliant, but I cannot recall any edited volume that could offer only first rate chapters. Nevertheless, in this book even the weaker chapters are valuable, providing information and ideas that are worthwhile to think about. Moreover, all the chapters are linked by a coherent theoretical theme: they all address the role networks of various kinds play in the postcommunist economic transition.

The book begins with a superb introduction by the editors. It not only elegantly frames the subsequent chapters, it also presents an important theoretical statement. Neoclassical economics always considered evolutionary selection another way of arguing that individuals are maximizers: nothing survives that is not optimal. Grabher and Stark argue that evolution would stop if only the fittest survived, unless, of course, the environment changed. In that case, however, there is no guarantee that the fittest are still the best suited to the new challenges. Counter to what neoclassical economics implies, what is optimal from an evolutionary perspective that always considers an entire population or system may look suboptimal at the individual level. They argue that solutions that proved to be inefficient yesterday must not be completely obliterated because they might be the answers to tomorrow's problems. To evolve, economic systems undergoing radical transformation need diversity, a wide repertoire of possible organizational forms, that includes not just winners but some of the also-rans as well. But Grabher and Stark do not stop at celebrating diversity. They try to identify how diversity should be organized so that the system can most efficiently take advantage of the information and opportunities diversity provides.

The introduction is followed by Stark's excellent article on recombinant property, originally published in an *AJS* article from which many of these theoretical insights have developed. Gerald McDermott, probably the best expert on this side of the Atlantic on Czech industrial organization, contributed a chapter on how privatization of the Czech industry brought about unintended and unforeseen consequences due to informational asymmetries and tight economic and social linkages. McDermott's analysis is rich in details while never losing sight of its theoretical polemic against the simplistic property rights theory on which Czech privatization

was built. Grabher's article presents the different strategies Western companies followed in investing in East Germany. Globally integrated new firms buy and sell on the world market and have few ties to local markets, except for labor. They either offer mostly low-skill jobs or take advantage of the traditionally high skills of workers, but culturally and socially atomize them, by cutting them off from their local communities. Either way, the region remains backward even if these companies prosper. Regionally integrated new businesses, on the other hand, rely on local suppliers and sell to local markets. They keep regional economies afloat, but their record of organizational and technological innovation is poor.

Judith Sedaitis, in a fascinating piece, analyzes the emerging commodity markets in Russia. Her thesis is that commodity markets will operate differently depending on the density of the social networks from which they emerge. Thus embeddedness becomes a variable rather than a theoretical assumption. When founding networks are of high density (have a high level of embeddedness) innovation is low, the supply of human and other resources is limited, legitimacy is high, and network members are trustworthy. Low-density founding networks, on the other hand, result in more competition, more creativity, a wider supply of resources and more opportunism.

The emerging small entrepreneurial sector is the topic of three chapters. Csaba Makó and Tibor Kuczi present a wonderful case study of a small enterprise network in Hungary that shows convincingly how social relations can be parlayed into formal and informal business associations and how business ties are reinforced by social ones. While Hungary started its small enterprise sector while still under socialism, Czechoslovakia had very few small entrepreneurs until 1989. Now on a per capita basis the Czech Republic has one of the biggest small private enterprise sectors in the developed world. Vladimír Benáček surveys these developments. He finds that small business cushioned the shocks of the transformation. He presents an informed if somewhat impressionistic account of the rise of the new entrepreneurs, distinguishing between operators, people who somehow had some connection with market forces before socialism, and the *nomenklatura*, who used their political power to get their start in private business.

István R. Gábor, who pioneered the concept of a small private enterprise sector under socialism, is more skeptical about the quick spread of small entrepreneurship. These enterprises remain very small, rarely growing into midsized companies. These microbusinesses suffer from all the disadvantages of smallness. Contrary to what many others have claimed, they are not hotbeds of capitalist mentality but rather the carriers of mentalities inherited from the socialist past.

Of the other papers, Peter Rutland's crisp overview of the politics of privatization and Wolfgang Siebel's history of the East German privatization agency, the Treuhand, deserve to be mentioned. This book is a major contribution to the debate on postcommunist economic transition, but it should be read by anyone interested in economic sociology.

Contemporary Capitalism: The Embeddedness of Institutions. Edited by J. Rogers Hollingsworth and Robert Boyer. New York: Cambridge University Press, 1997. Pp. xvi+493. \$59.95.

Alberto Martinelli
University of Milan

After the collapse of the Soviet Union, the spread of global capitalism, the fiscal crisis of the interventionist state, and the economic success of the "East Asian tigers," the thesis that free-market capitalism has won gained increasing popular acceptance. This view often implies that there is a unique model of economic growth and social development: that of self-adjusting market mechanisms.

This valuable and rich collection of essays challenges the assumption of the self-adjusting market on three major grounds: first, by arguing that economic activity is coordinated by several different institutional mechanisms and that a variety of capitalist models exist, as opposed to the idea of a single one; second, by stressing Karl Polanyi's notion of the embeddedness of economic institutions and, more specifically, the fact that markets and other coordinating mechanisms are shaped and shapers of social systems of production; third, by showing that specific forms of economic coordination are more likely to be used at some level of society (regional within a country, national, transnational regional, global) than at others.

The volume is structured around these three major lines of reasoning, plus a thoroughly informative introduction and a brilliant interpretative conclusion. In the first part, the chapters written by Boyer, Hage, and Alter, Coleman, and Sabel convincingly argue that no single institutional arrangement can pretend universal validity and that a more complex mix of context-specific and continuously evolving institutional arrangements coordinate economic activity—combining individual self-interest with social obligation and power with cooperation.

In the second part, the chapters written by Streek, Hirst, and Zeitlin, Coriat, and Hollingsworth, center around the concept of the "social system of production"—defined as "a configuration of complex institutions which include the internal structure of the firm along with the society's industrial relations, the training system, the relationships with competing firms, their suppliers and distributors, the structure of capital markets, the nature of state interventions, and finally the conceptions of social justice"; and examine the ongoing decline of mass standardized production in favor of flexible specialization, customized production, and diversified quality mass production. The authors differ with regard to the extent and the speed of the process of transformation, since some argue that flexible specialization will become the dominant social system of production—toward which various countries will converge—whereas others maintain that institutional inertia will place severe constraints upon the adoption of new principles.

The chapters of the third part by Grant, Hirst, and Thompson, Eden

and Hampson, and Schmitter, analyze another source of potential variation within capitalism, that of space. They explore the opportunities and threats of globalization on the performance of firms and of local, national, and supranational economies. Although all acknowledge the pervasive impact of globalization, the authors disagree on the role of regional free-trade blocs and the effectiveness of different national policies to adapt and control international competitive pressures.

In the conclusion, the editors propose a synthetic approach to institutional building that combines three major schools of thought: the new economic institutionalists (Williamson), the social embeddedness theorists (Granovetter), and the political science approach stressing competition for power and the role of the polity. It is, in a way, a revised version of Polanyi's typology of the three modes of transaction and appropriation: exchange, reciprocity, and redistribution. But it is updated and refined in order to give account of contemporary globalization. It centers on the concept of nestedness that, together with the concept of the social system of production, should pave the way for a theory of the emergence of new social systems of production and new world regimes. Nestedness, a complex intertwining of institutions at all levels of the world, derives from globalization and imposes the recombination of economic institutions at various spatial levels, since the embeddedness of economic institutions at the level of the nation-state has been progressively eroded both from above (supranationalization) and from below (regionalization). Nestedness implies that the evolution of capitalist institutions will produce a series of governance modes at various levels of society, and that economic policy and institutional change will be more difficult, since no supranational central authority is able to monitor effectively a series of innovations. The future of economic coordination is then uncertain, and the volume ends with a series of fundamental questions, the most basic of which is whether we have the capacity to govern ourselves democratically.

This rather lengthy summary has been necessary for a volume that is theoretically rich, complex, and stimulating and of great interest for scholars and students of economic relations from various social sciences. Its major achievement is to provide a coherent and multifaceted analysis of contemporary capitalism, by bringing together a valuable group of scholars (political scientists, industrial sociologists, institutional economists) with diverse backgrounds, paradigms, and perceptions of contemporary trends but sharing a common concern with the institutional context of economic activity.

The major limit of the volume is that theoretical richness is not matched by a comparable skill in formulating testable hypotheses and in assessing empirical evidence drawn from comparative analysis. Actually, sometimes improper conclusions are drawn from empirical evidence; it is true, for instance, that horizontal relations among firms such as networking, joint ventures, and alliances, have increased, but this does not mean that self-interest has given way to social obligation.

Besides, one has sometimes the impression that several authors do not

fully grasp the potential contribution of market mechanisms (competition, rational choice, pluralism) not only for economic efficiency, but also for social integration. There are plenty of arguments for stating the controversial role of the market in economic coordination. But the key idea is that only short-run and marginal choices can be left to the market and that the market must be tamed by alternative institutions, primarily political, in order to contain its explosive nature. The risk is that to the rhetoric of free marketers another rhetoric is opposed, that of trust and collective forms of governance, whereas the fundamental question remains how to achieve both innovation and consensus, both economic competitiveness and social cohesion.

Eco-Nationalism: Anti-Nuclear Activism and National Identity in Russia, Lithuania, and Ukraine. By Jane I. Dawson. Durham, N.C.: Duke University Press, 1996. Pp. xii+222. \$49.95 (cloth); \$16.95 (paper).

David R. Marples
University of Alberta

In the late 1980s, a series of unprecedented protests and demonstrations took place against nuclear power stations across the Soviet Union. Jane Dawson, an assistant professor of political science at the University of Oregon, explores the relationship between environmental activism and national independence movements using the example of three republics: Lithuania, Ukraine, and Russia, with some additional focus on Armenia and Tatarstan. Between 1990 and 1995, she carried out extensive fieldwork in the three major republics cited and monitored a persistent trend, namely the decline of environmental activism since the attainment of independence in 1991. This compact and well-researched volume seeks to explain why this was the case.

The book is divided into six chapters. Following an examination of social mobilization during the decline of communist regimes and in post-communist societies, the chapters look in turn at the different republics. Two chapters are devoted to Ukraine since the author includes an in-depth look at the protests against the Khmelnytsky nuclear station in detail. Chapter 6 focuses on the "nationalist enclaves" of Tatarstan and Crimea.

In Lithuania, the author notes the formation of the Movement for Perestrojka (*Sąjūdis*) in the summer of 1988. Furthermore, the outcry against the expansion of the nuclear station at Ignalina became a strong unifying factor for nationalist elements since the authorities did not initially characterize such protests as political by nature. Dawson perceives the protests as somewhat superficial, since they disappeared with independence. Moreover, many of the chief activists against the station in 1988–89 subsequently supported its continuing existence as the vital supplier of energy in the newly independent state. In Lithuania and Armenia,

environmental activism often served to camouflage national goals; these were less evident, she maintains, in Ukraine and Russia.

Ukraine is presented as a different case: a republic alarmed by Chernobyl into belated and tentative protests. Not everyone would agree with the author's assertion that national consciousness in the republic was weak because the history, languages, and cultures of Russia and Ukraine "are so closely intertwined" (p. 65). Much of western Ukraine has no historical links with Russia, while most observers would likely point out regional divisions in Ukraine as the key factor.

Similarly, the author's premise that there were no articles on nuclear power published in the Ukrainian literary press prior to 1986 (p. 68) is simply wrong. There were dozens, of which a few citations may suffice, all from the literary newspaper, *Literatura Ukraina* (see August 3 and 20, 1976; September 7 and 28, 1976; August 19, 1977; March 21, 1978; July 18, 1978; etc). There was a lengthy tradition of concern about nuclear power stations among the Ukrainian literary elite, which might have altered the author's conclusions had she been aware of the fact. Whereas she perceives antinuclear sentiment as developing belatedly because of a hardline Communist leadership until September 1989, it existed in passive form from the moment the first bricks of the Chernobyl station were laid.

Her conclusions on Ukraine remain valid, however, namely that the active *Zelenyi svit* (Green World) movement in Ukraine collapsed with independence since its anti-Moscow orientation was no longer necessary, the population was more concerned with daily subsistence, and nuclear power stations provided valuable energy for the struggling republic. In Russia, on the other hand, the protests were generally localized: the "not in my backyard" approach to nuclear power stations. Russians were unlikely to focus opposition on a Russian-dominated Soviet state. In Tatarstan and Crimea also, Dawson reveals that civic issues took priority over ethnic ones. The author thus concludes that although the environmental movement in the former Soviet Union was an important catalyst of a national struggle against the Moscow center, only in Lithuanian and Armenia were ethnic issues uppermost. The campaign dissipated with independence and, insofar as it exists today, it has taken on more refined forms of internal lobbying.

The volume contains some idiosyncrasies, not least the transliteration. Names are listed in the Russian form, but oddities abound. Thus one finds *Rabochaya Hazeta* (a mixture of the Russian version of the newspaper *Rabochaya Gazeta* and the Ukrainian, *Robitnycha Hazeta*). The director of the Chernobyl plant, an ethnic Russian, is mysteriously Ukrainized as Mykhaylo Umanets. The author also makes no distinction between the graphite-moderated reactors, such as Chernobyl, Ignalina, and Smolensk, now condemned as unsafe by the International Atomic Energy Agency, and other types of Soviet reactors. In itself, this distinction might assist in determining the credibility of environmental protests. Nonetheless, Dawson provides valuable documentation of the remarkable rise and fall of environmental activism in the former Soviet Union.

Russian Talk: Culture and Conversation during Perestroika. By Nancy Ries. Ithaca, N.Y.: Cornell University Press, 1997. Pp. ix+220.

Adriana Petryna
University of California, Berkeley

Cutting a different path through the by-now-traditional political and historical scholarship on perestroika, Nancy Ries offers a painstaking ethnographic portrayal of the moral and affective worlds governing that period. In *Russian Talk*, Ries considers how Russians, faced with social upheaval, collective fragmentation, and disorientation, commented upon the absurdity of their destinies and the injustice and suffering they experienced while negotiating everyday survival. Ries gathers narratives of lament, panic, cynicism, and relentless fatalism—voices of Moscow's intelligentsia and others—to make a major point: the very people who had the ability to evaluate the problems and to imagine other futures created obstacles toward "reasonable social evaluation and reformulation" (p. 17).

According to Ries's interpretive framework, talk shapes and maintains ideological and cultural stances, and constructs "models for life as much as models of it" (p. 3). Talk was not just a reaction to perestroika events, but it became "a significant constitutive aspect of the events themselves." Ries situates this talk within a distinct history that had a destiny of fading away in traditional historiography of that period, but that the ethnographer carefully assembled and analyzed as the very fabric of "social paradigms and practices" (p. 4). Such talk and characters in it reproduce what Ries insists on registering as "Russianness," but which can also be seen on the author's own terms as a painful and present amalgam of collective phantasms in light of uncertain futures.

While learning to hear in that transitional context, Ries was drawn into a question asked repeatedly by her colleagues: Why is Russian experience so full of suffering and misfortune? She suggests that the regular and everyday posing of this rhetorical question was itself "symptomatic" of "Russianness," helping to maintain the "kind of social and cultural institutions which perpetuate that 'suffering'" (p. 5). At this point, the reader should not expect to be drawn into a thorough discussion of local informal economies or into a genealogy of these regularly posed discursive practices as they relate to the concept of *kulturnost* or other Soviet collectivizations of subjectivity. Instead, Ries crafts a cultural history and classification of these discursive forms of suffering—laments, litanies, and mystical poverty narratives. She registers the key words, symbols, and scenarios employed in these narratives that, in her view, reproduced suffering by "discursively naturalizing" it. Ries explores stock figures in folklore and in real life that contributed to the efficacy of this conversational trance of sorts—the enduring female and the self-sacrificing grandmother (even the mischievous drunks were "transcending social orders"). Ries interprets the ways perestroika talk practices were employed as a form of Bourdieuan *hysteresis* of habitus ("the structural lag between opportuni-

ties and the dispositions to grasp them"; p. 188). Instead of "nourishing democratic structures and social relations" (p. 41), this Russian talk habitus stood still, dramatically reproducing social power and powerlessness (we the victimized *narod*, they the villain).

The affective content of these narratives of destinies of missed opportunities is not tragic, however. The reader, at times, senses their massive redundancy both in conversations and in the ethnographic text, but that content had a calculated effect. For the native listener, according to Ries, litanies and lamenting effected moral communities. In this world, Ries argues that suffering was currency: the primary gain as she experienced them was public entitlement to conceptualize social powerlessness and, paradoxically, to become endowed with a quasi-divine omnipotence of moral thought. There is no "talking" cure, and one cannot be talked out of whatever it is. In these worlds of discursive suffering, emerging market-driven social distinctions and novel values of personhood, wealth, and poverty were being acted out and realized. But again, they were only acted out in this world of Russian verbal appearances.

Between the lines, the reader is left wondering not so much who the ready-made powerless are, but what their lives appear like after the conversations are over. In which (forged, fractured, estranged, repetitive) webs of love, friendships, family relations, and informal economies—outside the anthropological one—do their sufferings transact? Upon which silences are their speech genres dependent? These points aside, the reader concludes *Russian Talk* with the sense that Ries has provocatively pointed to the directive powers of phantasmic presences in daily life. Today, in this Russian market transition, that habitus is muted and the currency of its mythical anatomy has been devalued. Yet its translation into new cultural sites with other values and forms of commodification remains a continuous challenge to be accounted for.

Rise and Demise: Comparing World-Systems. By Christopher Chase-Dunn and Thomas D. Hall. Boulder, Colo.: Westview Press. Pp. 322. \$60.00 (cloth); \$22.00 (paper).

David A. Smith
University of California, Irvine

No one can accuse these authors of a lack of ambition. Their task is no less than to explain the sweep of thousands of years of social evolution. Along the way they draw on an immense and diverse storehouse of existing literature from disparate fields of archaeology, history, political science, anthropology, and sociology, spotlighting several of today's burning scholarly controversies in macrostructural analysis. They also weave in interesting strands of their own recent research. A most impressive agenda.

Most comparative-historical sociologists working in the tradition of

Immanuel Wallerstein seek to understand the workings of today's capitalist world-system. Chase-Dunn and Hall are part of a group of interdisciplinary scholars eager to push much farther back, exploring how "many small-scale intersocietal networks have merged into a single global political economy" (p. 1) over the past 12,000 years. They claim that all patterns of intergroup interaction, even those of stateless and classless kin-based hunting and gathering peoples (which Wallerstein would call "minisystems"), can meaningfully be referred to as "world-systems." They redefine "world-system" to refer to the whole social context in which people live and the material networks important to their daily lives" (p. 28). Though that sounds amorphous, the authors specify four types of exchanges or interactions that bound their world-systems: (1) bulk-goods trade, (2) prestige-goods exchange, (3) political/military relations, and (4) information flow. To compare world-systems, the varying sizes of these networks and the degree to which they are distinct, overlapping, or nested are key empirical issues—as is the presence of core-periphery differentiation and hierarchy.

While they share an interest in long-term social change, Chase-Dunn and Hall reject the "continuationist" argument for 5,000 years of capital accumulation made by Andre Gunder Frank and Jonathan Friedmann. They acknowledge that limited commodity markets developed in the ancient world, but they concur with the widely accepted view that full-blown capitalism emerged in 16th-century Western Europe. Yet they insist it is wrong to see Europe as separate—instead it was a "subregion" of a much larger Eurasian "central system," which Frank and others claim developed hundreds of years earlier. The authors make interesting suggestions about the links between the advanced Chinese "core" and the European "periphery," mediated by the trading and raiding of central Asian nomads. It is intriguing that two relatively isolated islands at the opposite ends of this giant intersocietal network (England and Japan) developed similar feudal and urban systems, conducive to nascent capitalism, at about the same time.

The book's explicit goal is to understand the transformation of "modes of accumulation" from "kin-based" (based on "normative" social cohesion) to "tributary" (where "organized coercion of labor" predominates) to "capitalist" (based on "commodification of goods, land, wealth, and labor") and, hypothetically, to "socialist" (where popular democracy and "collective rationality" would predominate). Chase-Dunn and Hall also lay out a series of formal hypotheses and research strategies for comparing world-systems, suggesting a nomothetic variable-oriented approach to "testing" theories about trade networks, hierarchies, dominance, and the role of semiperipheries.

But the most interesting parts of the book use in-depth, historically rich data to illuminate issues of "transformation" of particular societies. A chapter on "new territories" (drawing on Hall's work on the historic frontier in the American Southwest) makes a strong argument that previous world-system definitions of "external arenas" fail to capture the extent

to which North American (or central Asian) nomadic tribes were actually "incorporated" (and forever changed by their contacts) at very early periods. City-states and frontier "marcher states" of the ancient world are depicted as frequent sources of innovation and change, evoking a generic image of the dynamism of the semiperiphery (though their definition of this concept remains a bit "fuzzy"). And the chapter on prehistoric northern California as "a very small world-system" linked to distant groups via "down-the-line trade" is a clever, well-researched (if not entirely convincing) attempt to apply concepts of intersocietal networks and hierarchies to a novel case.

Chase-Dunn and Hall have written a useful and interesting book, including the occasional pithy quote (e.g., "Nature is not a uniform plain upon which humans erect their condos"; p. 99). But it is not elegant and polished: the authors warn us that "the entire book is only the first round of an iterative process" (p. 8). Early chapters provide almost encyclopedic coverage of alternative approaches, however the Chase-Dunn/Hall argument sometimes gets submerged. An analysis of long-term quantitative data on city and empire size seems very preliminary. Finally, while the authors' grim assessment of the contemporary global situation (capitalism in crisis, the threat of world war and planetary annihilation) and call for a global solution (a democratic federated world state) are reasonable, the connection between these conclusions and the comparative study of the ancient Sumerian or Wintu tribal world-systems is rather remote. Maybe Chase-Dunn and Hall will make a more convincing case for that linkage in their next book.

Spaces of Globalization: Reasserting the Power of the Local. Edited by Kevin R. Cox. New York: Guilford Press, 1997. Pp. 280. \$44.95 (cloth); \$22.95 (paper).

Janet L. Abu-Lughod
New School for Social Research

A quarter of a century after scholars discovered the "new" globalization of the world's economy (an event signaled in part by the appearance of Fotker, Foekel, and Kreye's *New International Division of Labor* [Cambridge University Press, 1980]) much-needed correctives are being introduced—correctives that stress the residual importance of localities, national differences, and (still to an inadequate extent) history in positioning players in the increasingly global capitalist system. Earlier dramatic but perhaps premature generalizations about the communications revolution, the emergence of footloose transnational corporations, and the fungibility of high finance capital are being modified through concrete detailed research about difference. Even M. Castells, one of the most "convinced" of the new globalists (see his *The Informational City* [Blackwell, 1989]) has now begun to emphasize local (primarily national) variations and

contingencies related to the communications/network revolution. See his three-volume magnum opus, of which only the first two volumes have appeared: *The Rise of the Network Society* and *The Power of Identity* (Blackwell, 1996, 1997). The book under review is one of the most empirically grounded attempts to restore space to its central position in the discussion, an emphasis not unexpected from the (mostly) geographers and regional scientists who have contributed to it.

The underlying thesis of the volume is captured in the neologism "globalization," which, despite its ugliness, is probably with us to stay. This approach is most clearly put in Erik Swyngedouw's chapter, in which he rejects any antinomy of local and global: "The crux is not, therefore, whether the local or the global has theoretical and empirical priority in shaping the conditions of daily life, but rather how the local, global, and other relevant . . . geographical scale levels are the . . . products . . . of sociospatial change" (p. 140). Each of the articles explores this thesis in a different arena.

Michael Storper's masterly introduction sets the scene by arguing that "global capitalism is being constructed through interactions between flow economies and territorial economies" (p. 31). The chapter by Meric Gertler illustrates this distinction, demonstrating that although flows of finance capital may be fairly deterritorialized (a point later questioned), the actual production, distribution, and uses of machine tools are inevitably territorially grounded. Andrew Mair's chapter focusing on the reorganization of Honda offers a concrete demonstration of this proposition. Far from a monolithic transnational corporation producing a common "global car" for a generalized international market, Honda's production is now segmented territorially and its products are designed for localized outlets, a process he terms "strategic localization." Gordon Clark and Kevin O'Connor even question whether markets for financial instruments, usually viewed as an ideal-type of globalization, are either as centralized or as fungible as others claim. They distinguish three types of capital investments: "transparent" products, such as gold, where prices vary little on the global scale; "translucent" products, such as balanced equity products, which are traded within national markets; and "opaque" transaction-specific institutional financial products, which, because they require specialized expertise and entail high risk, are best traded locally. Kevin Cox sums up these three articles by redefining capital's relation to space as a variable, contingent upon product and strategic considerations. "Globalization is but one [strategic] possibility" (p. 132).

The collection then shifts to the related question of labor. Absolutely unique in the literature is a historical perspective and a conceptualization of labor not as a mere victim of global capital, but as an active agent beginning to operate globally as a counterweight to globalized capital. Chapter 7 by Andrew Herod fills this gap heroically, making us more aware of the historic role labor has played in the past century or more, when global capitalism was taking root. The historical lacuna is also partially filled by the next chapter by Ton Noterman, which traces the glob-

alization of capital back to the 19th century. (My private cavil with the literature on globalization is its tendency to start the narrative at 1973!) He supports his contention that "whereas international capital mobility has indeed increased dramatically during the last 25 years, its present level is comparable to the situation before the Great Depression . . . [with] the pre-1914 period . . . marked by very high capital mobility" (p. 203). He thus follows the line of reasoning David Gordon explored so convincingly almost 10 years ago in his "The Global Economy: New Edifice or Crumbling Foundations?" (*New Left Review* 168 [1988]: 24–65). His attempt to connect "capital vs labor" struggles over time to the possibilities for social democracy is expanded in the final chapter by Murray Low on democracy's prospects and how these have been obfuscated in the current debates over globalization.

This book, while scarcely "light" reading, thus makes a substantial contribution to a new paradigm about globalization. The volume is theoretically coherent, well documented, and mercifully free from errors. It would be petty to point out, therefore, that David Harvey's new book, *Justice, Nature and the Geography of Difference*, so relevant to the evolving paradigm, has been twice attributed (pp. 140, 163) to someone named Harvery. But the next edition can easily correct this. I suspect this gem will eventually be discovered and have a long shelf life. It certainly deserves careful attention.

Consuming the Romantic Utopia: Love and the Cultural Contradictions of Capitalism. By Eva Illouz. Berkeley and Los Angeles: University of California Press, 1997. Pp. xv+371. \$45.00 (cloth); \$17.95 (paper).

Scott Coltrane
University of California, Riverside

Sociologists have long been concerned with explicating the connections between theory, culture, and economy, but only in the past several decades have they seriously attempted to understand how love fits into the picture. Some of the most provocative of these attempts, including Frankfurt school critical theory, Birmingham cultural studies, and French social theory from Foucault to Bourdieu, have explored how emotions and cultural practices are shaped in the crucible of capitalism. In *Consuming the Romantic Utopia*, Eva Illouz provides a window on modern romantic sensibilities, contributes to ongoing debates about the meaning and practice of love, and speculates about the theoretical implications of the increasing commodification of romance. Relying on historical, literary, and sociological analyses, Illouz asks how specific vocabularies and images of love become publicly available. Her interest is in "bringing love within the traditional business of cultural sociology and submitting it to questions of which tacit meanings and symbols organize our romantic experience" (p. 5).

Illouz combines theoretical insights from divergent traditions, incorporating ideas from phenomenology and everyday life perspectives; she also draws on ideas about culture, society, and economy from classical theorists like Marx, Weber, and, especially, Durkheim. Eschewing the question of whether economy determines culture or vice versa, she focuses on how they are reciprocally constituted. Illouz shows how rituals of dating, lovemaking, and marriage are shaped by, and have shaped, popular images of leisure and consumption. Following Daniel Bell, she assumes that the culture of capitalism is self-contradictory and shows how romantic imagery and practice draws from conflicting cultural idioms of hedonism and work discipline. Romantic love under late capitalism is a liminal and utopian fantasy, conforms to media-produced stereotypes, is constituted through leisure consumption, and is achieved by rational, calculating individuals.

Although more questions are raised than answered in this book, the questions are important and the emergent understandings are delectable. Beginning with an examination of film and women's magazines, Illouz presents a historical account of the cultural portrayal of love in the first quarter of the 20th century. Her thesis, well supported by her data and by previous research, is that the theme of romance was used to promote a variety of products and increasingly became associated with consumption. She contrasts *candid* or direct consumption in advertisements with *oblique* consumption, in which romance is indirectly associated with leisure activities. By 1940, advertising and the movies had popularized "a vision of love as a utopia wherein marriage should be eternally exciting and romantic and could be if the couple participated in the realm of leisure" (p. 41). Photographic and cinematic realism helped to stereotype romantic vignettes, promoted identification with romantic heroes, and instructed teenagers in love. Advertising provided "a rich lexicon and a powerful syntax to formulate visions of harmony, intimacy and success" (p. 80). Through romantic imagery, commodities became associated with erotic attraction and achieved intimacy; in turn, the commodities themselves became objects of desire.

Relying on cultural texts such as magazine articles and popular songs, Illouz shows how a hedonistic-therapeutic model of marriage came to depend on a combination of pleasurable leisure consumption and hard work to keep the relationship romantic. As couples went from calling to dating and moved from parlor to car, romantic love became closely associated with a middle-class culture of respectability and "with forms of 'positional' consumption, aimed at displaying status in a competitive system of social stratification" (p. 71). Although she later claims that, in utopian romantic imagery, "gender divisions, social identities, and class inequalities are negated by the mutual access to leisure, the suspension of responsibilities, and the fusion of selves with nature" (p. 181), Illouz hints that social class differences are also constructed through romantic leisure consumption.

In the second half of the book, Illouz moves beyond textual and histori-

cal analysis to document how contemporary consumers, as experienced viewers of romantic intimacy in film and advertising, construct and discuss their love lives according to popular stereotypes. Gender differences are barely discussed, but interviews with 50 men and women reveal some interesting commonalities and differences across class lines. By asking her respondents to comment on their own romantic encounters, to answer questions about relationship vignettes, to evaluate photographs (plates reproduced in an appendix), and to judge greeting cards, Illouz shows how popular culture shapes personal narratives and helps construct romantic selves.

Some might challenge Illouz's suggestion that "the upper middle class is better equipped to cope with the cultural contradictions . . . of love" (p. 285) or reject the claim that gender differences are minimized in contemporary commercial imagery, but she has presented a forceful and compelling case that romance has become commodified and commodities have become romanticized. I would recommend this well-written book to any serious scholar interested in the sociology of love and will add it to the reading list for my graduate students specializing in cultural studies.

Work, Ritual, Biography: A Muslim Community in North India. By Deepak Mehta. Delhi: Oxford University Press, 1997. Pp. 283.

James Zetka

State University of New York, Albany

Deepak Mehta's *Work, Ritual, Biography* is an unconventional ethnography of Muslim weavers in North India. Mehta's focus is on the relationship between practice and discourse. Discourse and practice are not disjunctive and separate domains for this community; they come together in a coherent unity informed by Islamic cosmology. This thesis is developed through an examination of everyday and ritual cloth production, quilt making, the circumcision ritual, and the practices of a mystic healer.

With regard to weaving, Mehta explores the relationship between bodily techniques and the religious cosmology that infuses them with meaning. He argues that work techniques and communal traditions represent a unity. This becomes most transparent when Mehta focuses upon the handloom in chapter 4. Islamic cosmology, as expressed in the text referred to as the *Mufidul Mu 'minim*, classifies each component of the loom and infuses it with attribute and meaning. The bodily movements enacted when working the loom are identified with spoken prayers. Technique and word become one, as weavers' bodily motions on the loom become part of the ritual.

These prayers, according to Mehta, establish a multipolyphonic dialogue between the weaver and God, as well as between the weaver and specific members of his household or extended family. In everyday cloth

production, for example, only the male head has the right to recite the prayers initiating the production cycle and those used while the loom is working. In prayer the male head stakes claim within the household to control the loom. Mehta tells of an instance where a younger brother, with an eye to establishing his own independent hearth, staked claim to a second loom used in his familial hearth by reciting prayers while using it. The older brother, who was the head of this hearth, was forced unwittingly to relinquish control of this loom to the younger brother, who had imposed his signature and "word" upon it through prayer.

Mehta provides his readers with nuanced detail regarding this community's weaving practices and the traditions they embody. He describes the stages of the work process, how the household is organized spatially for cloth production along gender lines, how the kinship system assigns individuals to these processes on the basis of generational and genealogical position, how traditional weaving techniques structure the lifeworld and infuse meaning to time, space, and physical bodies themselves, how ritual production for the shroud is linked symbolically to the elements of nature, and how each of the loom's 15 components relate to valued attributes through prayers from the *Mufidul Mu 'minim*.

Chapter 5 provides an interesting contrast to the chapters on weaving. The focus here is upon quilt making, a work process that women control exclusively. Where cloth making constitutes a highly tradition-laden work process legitimated by religious authority, quilt making is not subject to the same types of scripted conventions and categories. Mehta discusses in detail the tools, techniques, and dialogue employed in making both everyday and embroidered quilts.

Mehta's analysis of the dialogue established between quilt maker and quilt is interesting. In everyday production, the quilt maker engages in a dialogic relation with her hearth members through the quilt within which she judges and comments upon their character on the basis of the rags she uses from them, a type of dialogue not legitimated in face-to-face interaction with males. She uses this dialogue to express her experiential situation in the hearth—one of pain and suffering.

Through stitching quilts, Mehta suggests that women code the memory of the hearth and of the extended family. He also argues that, in the collective production of embroidered quilts for the marriage ritual, the contingent nature of the work process enables women some license in introducing innovations into the tradition and in arguing for such innovations' appropriateness and legitimacy. Mehta draws provocative linkages throughout chapter 5 between the techniques of quilt making and women's peculiar social status within the affinal kinship system. The following chapters interpreting the circumcision ritual and the practices of the mystic healer are also interesting.

There is much that is commendable in *Work, Ritual, Biography*. Mehta's emphasis on the importance of how the body, space, and time are structured and on how each interrelate with one another and with tradition makes for an unusually detailed case presentation. While their evi-

dentiary bases are not always clear, Mehta's interpretations are often provocative. This monograph should be of interest to culturalists, especially those specializing in Islamic societies.

However, the analytic content of the book interacts primarily with theoretical debates from semiotics, and the peculiar way in which Mehta frames his findings makes for difficult reading. I found it hard to follow many of Mehta's arguments. Also, I would have preferred more discussion of the general theoretical significance of his interpretations. The book's influence within mainstream sociology will likely be limited.

Time for Life: The Surprising Ways Americans Use Their Time. By John P. Robinson and Geoffrey Godbey. University Park: Pennsylvania State University, 1997. Pp. xxii+367. \$24.95.

Gary Cross

Pennsylvania State University

This is a provocative book. It begins by challenging virtually all the common views about changes in time use since the mid-1960s: that Americans, especially women, are working more today; that men's contributions to housework has not increased much; that parents spend less time with their children; and that time devoted to the arts, religion, and civic organizations has decreased. Conclusions are based upon time diaries from 1965, 1975, and 1985 (from the Americans' Use of Time Project under the direction of John Robinson). The authors argue that these data provide far more accurate accounts of how people use time than the time estimates or anecdotal evidence used by more pessimistic researchers.

The validity of the data—especially from the 1965 diaries, which critics contend overrepresent employed people—remains controversial. And the absence of data for 1995 limits the contemporary relevance of the study. Aggregate figures on changing hours in paid employment (dropping from 46.5 to 39.7 for men and from 36.8 to 30.8 for women over the 30 years) disguise the impact of rising rates of early retirement for men and part-time work for women. The data also often confirm the pessimism of others—for example, employed women had five hours less free time than employed men in 1985 and, if men in the 1980s did more shopping than they did in the 1960s, 80% of child care remained in the hands of women in 1985, representing little change. Most important, most of the six to seven hours of decline per week in combined market and housework occurred between 1965 and 1975, with little change observed since. This may suggest less a trend toward increased free time than that a "Golden Age of capitalism" ended more than 20 years ago. In any case, with improvements in productivity, one would expect a greater reduction in work time than found here.

Still, their data show a rise in housework for men (by 4.2 hours per week). And much of what Robinson and Godbey report is fascinating

and very suggestive: that labor-saving appliances and cars seem to save little work time and that time devoted to physical fitness seems to have declined in the 1990s. So thorough and careful has been the gathering and presenting of this data that no one can dismiss this evidence without equal effort. So although the data will remain controversial, they will also stimulate much clarification of methodological issues and interpretation.

Among the more important conclusions is the rise in free time of older Americans (rising by 11.6% for those between 55 and 64 years of age). Indeed, the authors conclude that contemporary America "hoards huge amounts of free time to be spent only in the last 15-20 years of life" (p. 309). With Arlie Hochschild and Juliet Schor (*The Time Bind* [Holt, 1997]), they agree that the time famine has particularly affected 25-34-year-olds (who even by their "optimistic" data have seen an increase in free time of only 0.6%). This very much fits the popular image. But they are correct to emphasize that this is not the experience of older Americans, especially those without children. They note also that the perception of time constraint seems to be growing (at least until recently) and that this pressure is greatest among the 35-44-year-olds. They also agree with the common view that Americans have less free time than Europeans, especially for vacations. Most disturbing, they find that almost all of the time freed from work was consumed by television.

But underlying all the data is a message, part of a deeper debate about ideas. Robinson and Godbey are not really saying that we are a nation of whiners, unwilling to recognize that economic growth and technology is freeing us from the burdens of overwork and gender inequality, although this sentiment does pop up in the text. Instead of emphasizing structural changes in the economy and family as do Schor and Hochschild, they stress individual time use. They argue that we have "time for life" if we would assume the responsibility personally and socially to choose it. It may be hard to envision that change without a serious restructuring of jobs and family relations. But they are right to note that satisfying leisure time requires a serious reevaluation of personal values. Following the insights of Staffen Linder (*The Harried Leisure Class* [Columbia University Press, 1970]), Godbey and Robinson argue that, because we fill our time with more and more goods and higher and higher expectations, we feel that we have no time. Thus we are "clinging tighter and running faster. In the process, we are rushing past life" (p. 318). And this is something with which we can all agree.

Out of Place: Homeless Mobilizations, Subcities, and Contested Landscapes. By Talmadge Wright. Albany: State University of New York Press, 1997. Pp. xiv+408. \$60.50 (cloth); \$19.95 (paper).

Something Left to Lose: Personal Relations and Survival among New York's Homeless. By Gwendolyn A. Dordick. Philadelphia: Temple University Press, 1997. Pp. xi+220. \$59.50 (cloth); \$17.95 (paper).

Steven Worden
University of Arkansas

It is Talmadge Wright's ambition in his formidable book, *Out of Place: Homeless Mobilizations, Subcities, and Contested Landscapes*, to bring together the elements of cultural theory, social movement theory, urban studies, and qualitative analyses of homelessness into a coherent picture of contemporary urban homelessness and the ways in which the dispossessed resist domination. Applying Castoriadis's concept of the "dominant social imaginary," Wright delineates the contending conceptions of "homelessness" constructed by middle-class activists, policy makers, city planners, and city officials, pointing out the manner in which these conceptions have real implications for the homeless caught in a hypermodern capitalist city catering to the finance, insurance, real estate, and entertainment industries.

In his naturalistically based analysis, Wright carefully documents the emergence, growth, and eventual demise of two abbreviated social movement organizations (SMOs). One SMO developed out of the ties between students from Stanford University and San Jose State University and the homeless living along the Guadeloupe River. The second organization grew out of an alliance between homeless living in "refuse" spaces in Chicago, local artists, community activists, and the "Mad Housers," guerrilla architects from Atlanta. Both of these movement groups mobilized what was previously thought to be an isolated, alienated, passive, and disaffiliated population into protracted struggles with city authorities.

According to Wright, identification with sociophysical space, coupled with ongoing political battles, enabled hut residents to create both a collective identity of "family" and a socially created "place" as the homeless passed from nomadic to settled status. For Wright, the residents in these two hut communities drew upon "submerged networks" creating "resistant heterotopias" within whose confines some privacy, dignity, autonomy, and freedom could be possible. In fact, as the hut dwellers in Chicago acquired greater visibility with the media covering their political struggles, they expanded their networks out to churches, community organizations, colleges, and even giving talks at local schools and colleges. Eventually the huts were cleared away and the residents dispersed. However, Wright argues that not only did the hut dwellers gain a sense of self-worth and efficacy in organizing, but students who became involved in alliances with the hut dwellers also gained new knowledge and aware-

ness of not only charity and altruism, but also a greater appreciation of social justice and inequality.

In contrast to the often highly abstract and prophetic tone of Wright's study of homelessness, Gwendolyn A. Dordick in her remarkable ethnography, *Something Left to Lose: Personal Relations and Survival among New York's Homeless*, admits, with somewhat winning modesty, that her book contributes nothing to the debate about what makes individuals homeless. Instead, she continues, her work documents the consequences of homelessness, especially in terms of its producing relationships that are mobilized to socially construct shelter "where it was not meant to be."

Dordick spent 15 months researching the homeless in New York City, studying social relationships across a range of settings: a clique hanging out in a large bus station, a group of shanty dwellers squatting in a vacant lot, gang life in a huge armory, and a small group living in a small private rehabilitative shelter. Interesting, the book, written in the representational style of a realist/classical ethnographer, also explores at some length the perils, ironies, and incongruities of a tall, young, relatively privileged white woman entering into the field and studying the lives of homeless, disproportionately male, black, impoverished urban dwellers.

At the bus station, the homeless survive through panhandling, scavenging, and "conversating" in relatively stable cliques. Cooperation underpins these cliques, formed for safety, company, and sustenance, even creating a "bank" that can be drawn upon collectively for food. "Holding out," if detected, can result in violence. Insightfully, Dordick notes that the very intensity of the group bonds militates against the development of friendship and trust between individuals, privacy, and the accumulating of sufficient resources for one to get off the streets.

The shanty community, in contrast, is comprised mainly of hustlers, prostitutes, and street merchants who are nearly all involved in heavy substance abuse, appreciate their personal dyadic relationships, and value closable doors. Bartering, not sharing, is the norm. As Dordick observes, "at the Shanty, survival is about buying and selling." Kindness is seen as weakness, as in the case of "Sammy" who extorts fellow residents for electricity illegally tapped from a Consolidated Edison light pole. Because of preexisting ties and the relative safety and privacy afforded by the huts, the shanty community allows for a richer pattern of relationships based upon trust, friendship, love, and dependence than what could be sustained at the bus station. Although shot through with suspicion, resentment, and addiction, the ties endure. In the third setting, the "Armory," Dordick dramatically overcomes crew members, staff, and residents in order to study a huge shelter of "last resort." Crews control virtually every aspect of shelter life from placement of cots, food lines, to providing protection. Residents build relationships out of "juice" and violence. "Marriages" between men also serve to provide a quickly communicated normative pattern of exchange and obligation in a violent and hostile environment. Oddly enough, it is in the safe, clean, and well-ordered setting of a small church-run shelter that Dordick seems to have the most

misgivings. Oriented around the ideology of Alcoholics Anonymous, encouraging moral betterment and productive activity, the shelter emerges as an overly administered social world run by a priest, lay social workers, and volunteers. Residents enjoy some material benefits and security, but because of space, they enjoy little privacy. In this restrictive environment, residents fear getting evicted and forgo interresident relations, preferring to concentrate on "staying on the good side" of volunteers and staff and working the program. Dordick apparently misses the liveliness, color, and drama of the other settings, absent in this world run by, from her perspective, ill-trained and moralistic volunteers and misguided social workers who exaggerate the importance of staying drug and alcohol free.

In reviewing these works, one concludes that Wright's and Dordick's studies complement each other so well they ought to be sold in tandem. Wright provides a highly theoretical macrolevel framework within which to view urban homelessness; Dordick's work depicts the homeless that perhaps would not be good candidates for addressing schools and universities. Wright's work captures the dynamics of collectivities working together for social change; Dordick gives us the day-to-day ambivalence of love and suspicion within shanty communities. While Wright perhaps glosses over the extent of substance abuse in homeless communities, Dordick shows its role in everyday life. While Dordick critically views administrators, staff, and volunteers, Wright still sees a role for volunteers and other allies who could work with the homeless. It is unfortunate that the same compassion, complexity, and depth of understanding that researchers routinely grant the homeless cannot be extended more often by those who, day in and day out, stay to work with them.

The New Middle Class and the Remaking of the Central City. By David Ley. New York: Oxford University Press, 1997. Pp. vii+383. \$85.00.

Richard Harris

McMaster University

In recent years a number of social scientists have argued that place is important in the constitution of social life. To these writers, the concentrations of a particular social group not only reflect its distinctive character but may also be used by group members to define and identify themselves. David Ley, a social geographer at the University of British Columbia, has been active in these debates, and his latest book presents a persuasive case study in support of the general argument.

In *The New Middle Class*, Ley examines the gentrification of many inner-city areas in Canada over the past quarter century, viewing these as exemplary of an international trend. Concerned with the social rather than the physical dimensions of inner-city change he defines gentrification to include not only the renovation of existing dwellings but also the redevelopment of inner areas through the construction, especially, the build-

ing of coops and condominiums. His purpose is to assess the extent to which gentrification has occurred in Canada, to review and analyze the major causes of that gentrification, and to explore its social, political, and cultural dimensions. His argument is that gentrification not only reflected the growth of what he—following Daniel Bell and others—refers to as the “new middle class,” but that, by expressing and embodying certain values, it helped that group define itself as a political and cultural force.

One of the great merits of the book is that Ley draws skillfully upon a diverse array of sources. Relying chiefly on quantitative data and an analytical methodology, the first half of the book examines the extent, the geography, and the determinants of gentrification (chap. 2, 3, 4). He suggests that there are three theoretical frameworks within which gentrification has been understood: the postindustrial thesis, the passage from Fordism to post-Fordism, and “complex arguments around postmodernism” (p. 11). He relegates to secondary status the “rent gap” theory proposed by Neil Smith in “Toward a Theory of Gentrification” (*Journal of the American Planning Association* 45:538–48; see also Smith’s recent book, *The New Urban Frontier: Gentrification and the Revanchist City* [Routledge, 1996]). Ley’s analysis of the geographical dimensions of gentrification is perhaps the most thorough that has been undertaken for any country, and his conclusion that inner-city labor markets played a vital role is significant. Those who have written about gentrification have agreed that its main negative effect has been the displacement of low-income households, and the removal of cheap accommodations from the housing market. Ley notes this effect and documents the sometimes dramatic changes in inner-city house prices that underlie it, but he does not explore in detail the impact of gentrification upon low- and moderate-income households.

In the second half of the book Ley draws upon a range of sources, including electoral data, surveys, and newspapers, in order to explore the political and cultural dimensions of his subject. He establishes the association of the new middle class with the emergence of grassroots politics at the neighborhood, and even the block, scale in the late 1960s. Here he points out not only its class but also its gendered aspects, as women activists brought their particular experiences and concerns into the political arena. He then traces the impact of community politics in the wider movements for urban reform, in the 1970s, and the subsequent evolution of a culture of consumption and “conviviality.”

More than most who have written about gentrification, Ley views the trend historically, as a product of a particular period in time and as a process that itself changed as the new middle class evolved. From this vantage point, in the conclusion he responds to those, notably Larry Bourne, who have argued that gentrification has never been the dominant trend in North American cities and that it has by now virtually run its course (see Bourne’s “The Myth and Reality of Gentrification: A Commentary on Emerging Urban Forms” [*Urban Studies* 30:183–89]). Ley suggests that, despite some economic setbacks and the aging of the

baby boom generation, it is still an economic and cultural force. In this connection I found it curious that he did not make a more explicit and sustained connection between gentrification and the so-called "New Urbanism" (Ley uses the latter term but only to denote the cultural sensibility that developed in the late 1960s and the 1970s). Among planners, architects, and some developers, the New Urbanism has been hailed as the way to make suburbs more urbane. Its aesthetics are similar to those of gentrified neighborhoods, and I suspect that so too are its class and demographic dimensions. One might argue that the rise of the New Urbanism proves Ley's case: the cultural movement associated with gentrification has had a large impact, now extending into the outer "burbs." In the process, however, its links with the inner city are being broken. The success of gentrification may be leading to its demise as a distinct process. Ley does not consider this possibility or its future significance for the identity of the new middle class. Even so, he has written one of the most nuanced and complete accounts of gentrification yet published. It is indispensable reading for anyone interested in the subject.

On My Own: Korean Businesses and Race Relations in America. By In-Jin Yoon. Chicago: University of Chicago Press, 1997. Pp. vii+274. \$45.00.

The Korean American Dream: Immigrants and Small Business in New York City. By Kyeyoung Park. Ithaca, N.Y.: Cornell University Press, 1997. Pp. vii+228. \$39.95 (cloth); \$15.95 (paper).

Edward Taehan Chang
University of California, Riverside

Korean immigrant experiences and their entrepreneurial "successes" have received considerable scholarly attention over the past two decades. Two recent books, In-Jin Yoon's *On My Own: Korean Businesses and Race Relations in America* and Kyeyoung Park's *The Korean American Dream: Immigrants and Small Business in New York City*, are welcome additions to the growing literature on Korean immigrant entrepreneurship. Both books examine three key issues: immigration, entrepreneurship, and race relations. And it is interesting to compare and contrast how a sociologist and an anthropologist approach these issues.

Yoon's book is a historical and comparative study of Korean immigrant enterprise and the relations between Korean immigrant business owners and their non-Korean customers. In his book, Yoon makes extensive use of statistical data and tables and Korean-language newspapers. Park's book analyzes the impact of Korean small businesses on the formation of ideology, family relations, and gender-role restructuring. As an anthropologist, Park utilizes ethnography as a tool to conduct her research, relying heavily on personal interviews and participant observation. As such, Park focuses more on personal and family interactions and

emphasizes giving "voice" to Korean immigrant small businessmen, women, and others.

Korean immigrant entrepreneurship is the core subject of both Yoon's and Park's scholarly inquiries. Yoon's discussion of theories of immigrant entrepreneurship is succinct, clear, and effective. For example, Yoon provides concrete data on the intergroup differences in self-employment rates and applies them to the case of Korean immigrants. Yoon argues that Korean immigrant merchants are not playing the role of the middleman minority because Korean retailers heavily rely on Korean immigrant suppliers, rather than on white suppliers, as the middleman theory predicts. Furthermore, Yoon criticizes the middleman minority theory for its negative political implications for intergroup relations, specifically Korean-African-American relations.

While Yoon's statistical data is impressive, his singular reliance upon it tends to reduce the people he studies to numbers. This is where Park's book provides a strong counterpart. Park also discusses the proliferation of Korean-owned businesses in the United States, but she gives descriptive profiles of specific Korean businesses in New York. Her approach humanizes the subject by letting people speak about their own lives in the United States. In comparison to Yoon's treatment, however, Park's discussion of why and how Korean immigrants gravitate toward self-employment seems rather superficial and incomplete. Her theoretical chapter, for example, lacks depth and provides little discussion of the existing literature on immigrant entrepreneurship. Furthermore, I was troubled by the "old" and "dated" data she presents in the book to support her general thesis; although this book was published in 1997, the information she presents is mostly from the 1970s and the early 1980s.

The process of immigration from Korea to the United States is another common theme addressed in both books. What motivated Koreans to immigrate to the United States? Both authors agree that structural factors, such as dependent development, class inequality, industrialization, and urbanization in Korea, contributed to the migration of educated middle-class Koreans to the United States. Yoon argues, however, that the post-1985 Korean immigrants tend to be less educated and from working-class backgrounds as compared with pre-1985 Korean immigrants. Park also discusses the structural factors involved in Korean immigration, and again she focuses more on personal accounts of the migration process from her informant's own perspectives.

Race relations and the formation of racial ideology is the third common theme of the two volumes. Although Yoon provides a descriptive and informative summary of accounts of Korean-African-American relations in Chicago, New York, and Los Angeles, his discussion does not go beyond the description of facts, figures, and events. His treatment of these important issues lacks an analytical and theoretical framework that would enable him to examine critically the role of race and ethnicity in the context of economic and demographic restructuring in America. Park's thesis is that the overconcentration of Korean-owned businesses

has had a profound impact on the development of ideologies in both public and private spheres. In particular, Park argues that Korean businessmen have been influenced by dominant racial ideologies. What is somewhat disturbing about her presentation of this subject is that she does not qualify or place in the context comments of her Korean informants. As such, their statements regarding African-Americans and Latinos, presented as direct quotations, perpetuate stereotypical images of these groups.

I was impressed by both volumes, but also found both flawed in some respects. Each of the three major parts of Yoon's book represent solid scholarship and thorough research based on empirical data. Unfortunately, it is lacking a theoretical framework that pulls the three parts together into a cohesive whole. Park attempts to delineate the experiences of Korean immigrants from their own point of view, an admirable objective. On the whole, therefore, the study lacks a clear focus and some of her assertions could use more supporting empirical evidence. For example, Park argues that the status of Korean immigrant women has improved greatly as family relations shifted from male dominant to women centered. As she presents the stories of only a few Korean immigrant families, this assertion seems to generalize experiences of all Korean immigrant families. This points to a major shortcoming of her book: the absence of a discussion on the role of the Korean immigrant churches on the development and perpetuation of a patriarchal ideology, which I believe still is prevalent in familial relations among Korean immigrants.

My last minor criticism of Park's book focuses on her choice of the McCune-Reischaur romanization system for Korean words. While she wishes to present the authentic voices of her informants, many of the transliterations are not readable or understandable to native speakers of Korean. Consequently, I found it extremely difficult to read, understand, and comprehend Korean words and their meanings.

I was particularly puzzled by Park's use of *anjong*. Park utilizes *anjong* as a metaphor to explain Korean immigrant experiences in America. According to Park, "*Anjong* is the key to understanding the Korean American dream, and provides a way to interpret particular social practices in the immigrant community" (p. 3). Park translates *anjong* as "establishment," "stability," or "security." One of the distinct and unique characteristics of Korean immigrants is their "risk-taking" behavior. For example, Korean immigrants are the most dispersed Asian group in the United States (they do not settle solely in areas where Koreans are concentrated). And, Korean immigrants in urban areas have often established their businesses in areas with high crimes. These tendencies reveal their willingness to take risks in order to realize the American dream quickly. As such, can *anjong* adequately account for risk-taking behavior? What is more, while *anjong* does have all the meanings she lists, the term generally means "stability." However, *anjong* is not a word Korean immigrants themselves use with particular meaningfulness and therefore Park's metaphor seems misplaced.

Despite these shortcomings, I highly recommend these two books for scholars, students, and anyone who is interested in Korean American issues and race relations. These two books complement each other in their different disciplinary approaches and emphases. I am certain that both studies will generate further discussions and debates about the role of Korean immigrant entrepreneurship in American life and its impact on the development of Korean-American ideology and consciousness.

Just Another Car Factory? Lean Production and Its Discontents. By James Rinehart, Christopher Huxley, and David Robertson. Ithaca, N.Y.: Cornell University Press, 1997. Pp. xi+249. \$37.50 (cloth); \$14.95 (paper).

Richard Florida
Carnegie Mellon University

Just Another Car Factory is a welcome and useful addition to the growing body of research on Japanese automotive factories in North America. Spurred by a wave of Japanese investment beginning in the mid-1980s, these "transplant" organizations captured the attention of social scientists interested in the transfer of Japanese-style management abroad, the adoption of innovative "high-performance" work systems, and the more general question of the shift from traditional "fordist" mass production to new and more advanced modes of economic organization.

The book focuses on the experiences of workers at the CAMI plant in Canada—a joint venture between General Motors and Suzuki—opened in 1986. It provides a comprehensive and highly readable account of factory life based upon a wealth of empirical information culled from shop-floor observation, personal interviews with managers, union officials, and factory workers, and survey research on the adaptation and commitment of shop-floor workers. The authors—two Canadian sociologists and a representative of the Canadian Auto Workers (CAW)—use this information to advance two interrelated arguments.

The first argument is steeped solidly in their empirical case. The authors basically contend that the CAMI factory is similar in fundamental respects to the traditional mass-production factories operated by the Big Three. They muster considerable evidence from interviews and surveys to show how worker commitment to the plant declined steadily over time, as did worker involvement in participative activities, leading ultimately to a major strike—the first at a transplant factory—in 1992.

The authors move from this empirically based argument to advance a second, more critical, thesis. Taking aim at the generally accepted view that transplants are at the cutting edge of the shift from traditional mass production to a new and more efficient system of lean production, they contend that the problems experienced at the CAMI plant stem fundamentally from the limits, tensions, and internal contradictions of lean pro-

duction systems. They go on to argue that lean production essentially represents a more insidious form of mass production, designed to pump more work out of workers through a combination of a fast-paced work and peer-enforced discipline. They take issue with those who suggest that lean production has transcended mass production, arguing that such a view is misguided and premature.

This second argument, while certainly provocative, is not persuasive. While it is always problematic to generalize from a single case, the CAMI case is particularly challenging because the plant is unique among the broader population of transplant factories. CAMI is a joint venture between General Motors—the world's largest automaker and a company that has long exemplified many of the core features of mass production—and the much smaller Japanese automaker, Suzuki—a company that is not known for its commitment for advanced work and production organization. Certainly, no one would argue that Suzuki represents either the workplace innovation or the economic power of a Toyota, the company which is most closely associated with lean production. Furthermore, the plant's workforce is represented by the CAW, which has actively opposed team-based work and other features of advanced work systems.

While the authors' effectively make the case that CAMI is neither an advanced work environment nor a workers' paradise, they are ultimately unconvincing in their attempt to generalize from these specific facts to a sweeping indictment of the transplants and advanced production systems. It is especially difficult for the authors to score because their work flies in the face of numerous case studies of other transplants, of careful plant-level comparisons such as those carried out by the International Motor Vehicle Program at MIT, and of several survey-based studies of both Japanese transplants and of U.S.-owned manufacturing plants.

The authors should be commended for the quality of their empirical research and for highlighting the negative conditions that can occur in specific transplants. While CAMI may well be "just another car factory," even such a well-documented case study will fail to convince a broad and knowledgeable audience that there is nothing new in the transplants or in the lean production system.

Fate and Fortune in Rural China: Social Organization and Population Behavior in Liaoning, 1774–1873. By James Z. Lee and Cameron D. Campbell. Cambridge: Cambridge University Press, 1997. Pp. xxi+280. \$49.95.

William Lavelly
University of Washington

This is a landmark work on Chinese social history. It pioneers the banner registers, a new source of data on late imperial China, and it raises Chinese population history to a new level of methodological sophistication.

It offers an unprecedented view of social stratification and mobility at the lower end of the social spectrum in premodern China. And it presents very intriguing data on demographic processes that are unlike any other data available for premodern China. The banner registers consist of triennial household censuses of rural military settlements. For one such settlement, Daoyi, a banner village of roughly 3,000 people, the authors have linked individuals, reconstituted households, and constructed vital rates from 26 enumerations over the course of a century. No previous work of Chinese historical demography has followed a commoner population for so long or in such detail.

Lee and Campbell set out the thesis that the life chances of the common man are shaped by two contradictory principles. On the one hand, position in the extended family hierarchy is an important determinant of social and economic success. On the other hand, individual attributes offer the possibility of rising in an institutional hierarchy beyond the household. They support the thesis using demographic behavior as a proxy for success. They propose that early marriage, many children, and long life are useful indicators of success in an agrarian society (although they quickly concede that death rates have little relationship to status). Using techniques of event-history analysis, they argue that the low fertility of men prior to attaining status occupations indicates that selection into these statuses are open to men of talent; high fertility after attaining status demonstrates that status confers demographic advantage.

The idea of "reproductive success" as the ultimate mark of status would appear to be borrowed from evolutionary biology, but Lee and Campbell steer clear of this association. Only at one point, when seeking to explain the higher death rates among the fertile male elite, do they seem to evoke the evolutionary paradigm, speculating about the "cost of fertility." However, the costs of fertility are generally born by the female of the species, and their data show no evidence that fertile females suffer higher death rates. In all likelihood, lifestyle factors (alcohol consumption?) account for the elevated mortality of elite males.

Lee and Campbell also press the thesis that Chinese families actively planned their families. One of the most stunning findings of the book (paralleled in Lee's other work on the genealogy of the Qing Imperial Clan) is that female infanticide must have been used on a fairly routine basis for the tailoring of family size and composition. The authors estimate that Liaoning peasants eliminated one-fifth to one-quarter of infant daughters, a scale that, if it is generalizable, has truly important implications for Chinese population dynamics in premodern times. But family planning was not limited to infanticide. The authors argue that Liaoning peasants controlled fertility as well and that they did so consciously and rationally. There is evidence for this, but there are also problems with this evidence. Liaoning peasants, as in other Chinese populations, had very low marital fertility, "so low that they are almost inconceivable without assuming widespread sexual restraint and/or the technology for family limitation" (p. 92). But restraint and technology imply purposeful fam-

ily limitation, and it is quite conceivable that other customs—long-term breast feeding, for example, or customs affecting frequency of intercourse that did not have fertility limitation as their purpose—could account for long intervals between births.

The authors speak of "fertility decision making" (p. 101) and note that "fertility decisions . . . were sophisticated and deliberate" (p. 102). But the evidence for this proposition is more ambiguous than their exuberant language suggests. Certainly the fact that reproduction varies by social class is not evidence of intentionality (this is implied in p. 96, n. 22, which states that "couples adjusted their fertility" to produce this variation). The authors demonstrate convincingly that Daoyi fertility responded to economic conditions. However, restraint of fertility is strongest for married couples who had not yet begun childbearing (p. 100). This fact appears inconsistent with purposeful control; rather, it would seem as likely that bad economic conditions influenced the living arrangements of couples early in marriage, perhaps increasing the incidence of "delayed transfer" or posing some other involuntary impediment to the birth of a first child.

As a tightly controlled military colony, essentially closed to migration and subject to regulations affecting land and family, the banner community is a special organizational environment. This environment might account, for example, for the radical evolution of household structure observed over the course of the century. For example, households containing cousins or uncles ("diagonal households") were virtually nonexistent in 1792 but were the norm a century later. The authors cite long-term increases in Malthusian pressures to account for fundamentally altered demographic and social behavior. These pressures were seemingly created by banner containment. What we have, then, is a study of demography and social mobility under a special (although, as the authors argue, common enough in northeast China) set of conditions. As with Arthur Wolf's classic study of a deviant marriage form that flourished in Taiwan, the value of the work lies not in its "typicality" but in what it can tell us about the basic tendency, and plasticity, of social behavior under the great variety of conditions that existed in China.

How to Succeed in School without Really Learning: The Credentials Race in American Education. By David F. Labaree. New Haven, Conn.: Yale University Press, 1997. Pp. vii+323.

Christopher J. Hurn
University of Massachusetts

David Labaree is a fine educational historian and Yale a distinguished press, but this book is a case of "caveat emptor." First, although the title and subtitle promise an argument about learning and credentials, and, by implication, evidence of a negative relationship between the two, no such argument is presented. Indeed, only two or three pages are directly

concerned with the topic. In the introduction Labaree states that he will "show that the process of getting ahead often interferes with getting an education, and that the process of getting an education frequently makes it harder to get ahead" (p. 1); but no such demonstration of either of these points is attempted in what follows. Second, nowhere in the jacket material, the acknowledgments, the introduction, or even the first 263 pages of the book, is there an indication that this is not an original manuscript; instead, with the exception of one chapter and part of another, this is a collection of previously published essays. Nor is there any one place in the book when the reader can find this out—only by reading all 10 prefaces to the end notes of each chapter can the reader judge the extent of prior publication. Many readers, I suspect, like this reviewer, will be misled—discovering that this book is not what it purports to be only after a struggle to find a sustaining argument connecting such disparate topics as "The Carnegie Cult of Social Efficiency," "Rethinking the Movement to Professionalize Teaching," "The Rise of the Community College," and "The Lowly Status of Education Schools." Had the author told us clearly that the chapters were written for different occasions and for different purposes, the absence of a coherent sustaining argument would have been less bewildering.

All this is rather a pity because many of the individual essays are fine enough to deserve reprinting. In all the chapters mentioned above the historical narrative is beautifully handled and the writing has a nice, critical edge. Everyone concerned with teacher education should read the essays on the history of attempts to professionalize teaching and the lowly status of education schools. And if these pieces have a pessimistic or even elegiac tone, it is because the author demonstrates well how a pattern of missed opportunities will not change in the near future: Schools of education, precisely at a time when there is finally real knowledge available to help teachers teach, are now going full tilt away from a focus on teacher education in a vain effort to raise their status by emulating faculty in research universities. The one entirely new chapter, tracing the history of student promotion from the 1830s to the present, is also valuable. It concludes with the interesting speculation that recent challenges to the benefits of tracking, combined with the growing failure of educational credentials to equip students for the job market, has led to a sharp decrease in social promotion.

Unfortunately, the weakest chapter in the book is the first. "Public Schools for Private Advantage" traces the changing fortunes of three educational goals, social mobility, social efficiency, and democratic equality, from the founding of the Republic to the present. Labaree argues that in recent years the first of these goals has become dominant with what he believes are disastrous consequences: "undermining learning, reinforcing stratification and promoting a futile and wasteful race to attain devalued credentials" (p. 52) But Labaree does not explain why such a race *should* undermine learning, nor does he consider the possibility that a tightening connection between schools and jobs might encourage students to work

harder. Perhaps a credentials race heightens stratification, but perhaps a system of second and third chances (unlike a Japanese-style examination system) favors the mobility of talented low-status youngsters. The possibility is not considered. The other problem with this chapter (and with the short conclusion) is that the characters in the struggle for educational control are conflicting educational goals rather than human actors. The social mobility goal he argues implies a consumer perspective on schools, which in turn, demands that schools take the form of a "graded hierarchy" promoting "the stratification of education" (p. 28). But this confuses philosophy with sociology. One cannot demonstrate that a particular educational objective has certain consequences merely by teasing out the logical implications of that goal. Perhaps the key actors see these implications, but perhaps they choose to ignore them, especially when the goal (like "opportunity") is primarily rhetorical in the first place.

What we have then, is a collection of seven fine, though loosely connected, historical essays, one of the best of which is new and is preceded and followed by two chapters of much lower quality. My own sense is that these two more general chapters detract from the rest and that the whole is thus rather less than the sum of its parts. Other readers (and acquisitions editors in libraries) may prefer the convenience of this collection and ignore or discount its misleading presentation as a coherent book-length narrative of original scholarship.

A History of Young People in the West. Vol. 2, *Stormy Evolution to Modern Times*. Edited by Giovanni Levi and Jean-Claude Schmitt. Cambridge, Mass.: Harvard University Press, 1997. Pp. vi+409. \$35.00.

Susan B. Whitney
Carleton University

Historians have played an important role in the increased scholarly attention to youth during the past 30 years. They have highlighted the fact that "youth" is a socially and culturally constructed category, one whose meanings vary across time and according to class, race, and gender, and have drawn attention to the ways debates about youth have often been infused with a range of symbolic meanings. But if important histories of youth and generational thinking have been published, few collected volumes tracing the broad outlines of the history of Western youth have appeared. The publication of the two-volume *A History of Young People in the West*, which contains translated essays by prominent European historians, marks an important, if not entirely successful, attempt to do this.

Stormy Evolution to Modern Times consists of nine essays focusing primarily on the period between 1789 and 1945. Written by historians and art historians, the essays employ a range of methodological approaches (excluding those drawn from sociology, however) to address cen-

tral questions in the history of modern youth, including the ways youth has been defined and youth consciousness constituted, the interaction between young people and the institutions of the modern nation-state, the political manipulation of ideas about youth by 20th-century fascist regimes, and the multiple meanings and symbolic significances of debates about youth. Thus, for example, Daniel Fabre approaches the history of village youth through an anthropologically based analysis of a youth festival in a small village in the Languedoc in France, while Michelle Perrot draws on approaches from women's and labor history to examine worker youth in the 19th century. Jean-Claude Caron focuses on middle and high school education to analyze the new politically charged relationship between youth and education characteristic of the 18th and 19th centuries, while Sabine Loggia traces the transformation of the military experience in the modern era into something that became a pivotal part of the life cycle and a rite of passage for European young men. As she and other authors point out, the linking of military service to modern notions of male citizenship meant that military experience resulted in important instances of age-based collective identification. The essays dealing with 20th-century material focus primarily on the crucial relationships between young people, representations of youth, and fascist regimes. The art historian Laura Malvano analyzes the ways Italian fascists used an array of images of youth to symbolize Italian fascism, while the art historian Eric Michaud examines the Nazi attempt to transform young people into soldiers of the new Nazi idea. The volume ends with an essay by Luisa Passerini analyzing the discourses on youth in fascist Italy and in the United States during the 1950s.

Despite the range of approaches taken and the centrality of the themes addressed, the volume is less comprehensive than its title suggests. Although it purports to be "a history of young people in the west" (with the qualifier "in the west" absent from the cover), its coverage is in fact rather limited. Many of the 19th-century essays, for example, take France as their focus and are comparative only in the sense that they include occasional examples from other Western European countries. The 20th-century essays start and stop with fascism, with the exception of Passerini's contribution. Because of this limited geographic and chronological scope, there is no treatment of such important—and well-studied—topics as North American youth (outside the 1950s), the relationship between youth and communist regimes, or the emergence of new types of youth cultures in the post-World War II West. Finally, the volume suffers from having no introduction or conclusion. Although volume 1 includes a good introduction to both conceptual issues in youth history and to the collection itself, readers of the second volume are left to draw together the main themes and defining characteristics of the modern era on their own.

But if the volume does not deliver as much as its title and, often, chapter titles claim, some of the essays are very good indeed. One of the highlights is the essay by Michelle Perrot, which employs a variety of methodological approaches and mines a wealth of secondary and primary

materials to examine the many facets of the history of working youth (read: French working youth) in the 19th century. Perrot delineates the ways youth was defined for working-class young people; examines the ways working youth was represented and imagined by authorities, investigators, and social commentators; outlines working youth's relationship to institutions such as the military and the patriarchal working-class family; and describes young people's experiences at the workplace and in the city. Attentive to gender throughout the essay, Perrot insightfully compares the ways experiences, representations, and possibilities varied from young men to young women. Although not all essays are this successful, this is an important collection, one that should be of considerable value to those interested in the history of youth, especially Western European youth, in the last two centuries.

Moral Culture. By Keith Tester. London: Sage, 1997. Pp. 164. \$23.95.

Philip Smith
University of Queensland

In *Moral Culture* Keith Tester explores the shifting nature of morality in the contemporary, mass-mediated world. At the heart of the text are four rather loosely associated studies on indifference, duty, guilt, and virtue. While the first two of these chapters sit quite well together with their common discussion of media, suffering, and the self, the other two read more like stand-alone essays than chapters in a tightly integrated book. In looking at indifference, Tester draws on Alfred Weber and Anthony Giddens to point toward the globalizing tendencies of the mass media and to suggest that these have disturbing implications for individual subjectivity. Tester suggests that we experience distant wars and atrocities as "completely and utterly banal," "exceedingly uninteresting," and as "completely and utterly boring" (p. 28). Drawing upon Simmel's writings on the city and money, Tester explains this "blasé attitude" (pp. 32ff.) in terms of habituation to the presence of the media and television in Western life and in the need of modern individuals to maintain a sense of "reserve" or distance from their environment.

The discussion of duty looks to Hannah Arendt and David Riesman in exploring the modalities of morally responsible human subjectivity in the face of suffering (inner directed, tradition directed, Kantian, etc.). While Tester identifies a number of such forms he makes it clear that other-directed forms driven by the mass media and contemporary urban life have grown in influence while the Kantian categorical imperative has diminished in power. This "space of mediated appearance" (p. 75) reduces the human condition of life to a spectacle that is no longer subjectively compelling. As a result the cosmopolitan Western television consumer practices an indifferent objectifying gaze that moves fleetingly from crisis to crisis.

Tester's discussion of guilt revolves around the 1968 My Lai massacre and asks how ordinary young men could have become perpetrators of atrocity. The answer, we are told, lies in Nietzsche's concept of resentment and its role in promoting the externalized aggression of the pariah group. This is the least satisfactory chapter in the text. The links between the case study and the rest of the book are not very clearly developed, while Tester's explanation seems rather speculative and impressionistic. The final substantive chapter on virtue probes links between culture and morality through insightful readings of Gertrude Himmelfarb, Norbert Elias, and Zygmunt Bauman. Here Tester develops a telling analysis of these writers in terms of their orientations toward the projects of civilization (pragmatic civility), culture (expressive solidarity and humanism), and *Kultur* (exclusionary and elitist forms of cultural identity). While these distinctions are a very useful contribution to social theory, relatively little is done to join them to the themes of the earlier two chapters.

Tester is a powerful and provocative thinker, and *Moral Culture* will repay reading by those interested in the links between the self, the media, and morality in contemporary society. The text makes full advantage of Tester's formidable erudition, although there are some errors of fact and consequent interpretation. George Orwell did not enlist in the International Brigade as Tester claims (p. 6), but in the Spanish Militia, demonstrating a qualitatively different form of solidarity. The reason Conrad's novel *Lord Jim* does not "pay any attention to the fate of the drowned pilgrims" (p. 59) is that they did not drown (the *Patna* was towed unharmed to Aden by a French gunboat). So Jim abandoned his duty in the face of imagined or potential, not real, suffering. And Jim's profound sense of shame and his desire to redeem his public reputation indicate he is as much other directed as inner directed (p. 58).

Empirically minded readers might find a more enduring problem with *Moral Culture* to be an evidentiary one. Tester offers profound theoretical analyses of audience apathy, but begs the question: Are audiences really apathetic? In making this claim he depends to a large extent on brute assertion, his own experiences, and citations to other theorists. He might have done better to have looked at an established literature (e.g., focus-group audience ethnography) where sociologists have actually talked to people about their responses to news images of suffering and brutality. Published research documents a variety of attitudes and emotions and suggests that audiences can exert pressure via the public sphere. Consider, for example, the impact of media coverage of the Kurdish situation at the end of the Gulf War, or African-American (and many white) responses to footage of the Rodney King beating. Boredom, I would suggest, is one among several possible audience attitudes toward atrocity. It is the task of sociological inquiry to identify divergent readings of media texts and locate their origins rather than to assume a universal response. Such research, documenting diversity, engagement, and even occasional solidarity, could provide a resource for heading off the undercurrents of despairing fatalism (e.g., pp. 20, 143) and bleak technological and envi-

ronmental determinism (chap. 1) that haunt Tester's analysis of the cosmopolitan gaze.

The Magic of the State. By Michael Taussig. New York: Routledge, 1997. Pp. ix+206. \$59.95 (cloth); \$17.95 (paper).

Sally Falk Moore
Harvard University

Michael Taussig is an anthropologist who has long been dedicated to elaborating on Marx as a philosopher of the human imagination. Again and again Taussig sees himself as extending the "poetics of the commodity" (*The Nervous System* [Routledge, 1992], p. 4). He plays with variations on the familiar Marxist theme that things which encapsulate the human labor that went into their production are normally conceived of in a capitalist system as entities of given value in an abstracted system of comparative values. Reading this in reverse, labor and people can also become thing-like, and mystically, as in religious objects and fetishes, things can come to have the qualities of persons and to be saturated with power.

In this book, purportedly but obliquely about the idea of the state, Taussig has focused on the shifting, transformational quality of powerful meanings and their remarkable characteristic of being able to be lodged in many and various places, images, and hosts. Through the imaginary of supposedly fictional individuals seeking spirit possession at a magic mountain, he shows the way in which the imagery of violence and death are combined in imagined national histories (mysteries?) with triumphal moments, healing powers, public monuments, official documents, and the other flotsam and jetsam of remembered collective and individual experience. This Taussig combines with the puzzles of consciousness and of not knowing, and he adds the mystery of evoking spirits and invoking possession through ritual.

This is no conventional social science essay on politics. Taussig's style is surrealist. We are in the realm of Franz Kafka. He tells many tales and sketches many colorful pictures. Not everyone will have the patience for this raconteurism, nor the taste for it, but if you can float along on the flood torrent of stories, jumbles of word pictures, and asides—some learned, some sexual, some horrifying—there are rich commentaries embedded in all of this. His strategy seems to be to prepare the reader for the intellectual argument by engaging him or her through emotionally weighted, seductive, and assaultive prose.

By means of this mélange Taussig shows his way of approaching the symbolic realm. He attacks the neo-Durkheimian interpretations of many anthropologists as simplistic, formulaic codifications. In an earlier book, *The Devil and Commodity Fetishism in South America* (University of North Carolina Press, 1980, p. 9), he said, "Unless we also realize that the social relations symbolized in things are themselves distorted and self-

concealing ideological constructs, all we will have achieved is the substitution of a naive mechanical materialism by an equally naive objective idealism (symbolic analysis). . . . Unless we realize that the social relations thus signified are themselves signs and social constructs defined by categories of thought that are also the product of society and history, we remain victims of and apologists for the semiotic that we are seeking to understand." Again and again he returns to Marx. In his more recent book, *The Nervous System*, Taussig makes the point that "the everyday is a question not of universal semiotics but of capitalist mimetics," and he goes on to knock anthropological analysis by speaking of "the need to critique what I take to be a dominant critical practice which could be called the 'allegorizing' mode of reading ideology into events and artifacts, cockfights and carnivals, advertisements and film, private and public spaces, in which the surface phenomenon, as in allegory, stands as a cipher for uncovering horizon after horizon of otherwise obscure systems of meanings. This is not merely to argue that such a mode of analysis is simple-minded in its search for 'codes' and manipulative because it superimposes meaning on 'the native's point of view.' Rather, as I now understand this practice of reading, its very understanding of 'meaning' is uncongenial" (1992, p. 147).

I quote these earlier works because there the arguments are lodged in somewhat more conventional essayistic presentations than can be found in the present book. This one is flying free. *The Magic of the State* attempts, through a surreal, evocative, and unconventional form of dramatic writing, to give an immediacy to all the images, ideas, and powerful emotions that can surround the idea of the state and the state of possession. The reader is to *feel* what it is like to witness, or even to experience, spirit possession. The magic mountain to which people go on pilgrimage-like quests to be possessed by dead spirits is both dangerous and empowering. Many of the dead spirits being recalled played a role in the national history of their country. "Here we should be mindful of the fact that for most people for most of world history, spirit possession was the norm. . . . Such theatrics requires a stage in space and time—as in the case of this ex-colony with its death-space built from stately poetics of anti-colonial warring fused to apocalyptic triumphalism. The spirit queen's mountain is the supreme attempt to state and restage just this moving space" (p. 79). This book is a concentrated effort to do the same in prose.

The Magic of the State seems to have been developed in a seminar in the Department of Performance Studies at New York University in 1987. Events are frequently reported as if they had been the firsthand experience of the author, but the format disowns them as such. The clear allusions to the South American locale of Taussig's fieldwork and the spooky accompanying photographs and drawings give the whole an eerie verisimilitude. He seems to be saying, "You think this is fantasy, but it is and is not." Taussig's style of truth-telling roams in the fantasists privileged territory. On the slopes of the magic mountain, he deposits, here and there, a commentary on meaning and an acerbic critique of his colleagues.

Authority: Construction and Corrosion. By Bruce Lincoln. Chicago: University of Chicago Press, 1994. pp. xii+228. \$22.50.

Joel Kuipers

George Washington University

In Nevada on April 13, 1992, an antinuclear activist named Rick Springer walked calmly up to a podium where Ronald Reagan was speaking, pushed the president aside, and smashed a small gift that had just been presented to Mr. Reagan. He then took the microphone and began by saying "Excuse me Mr. President . . ." when he was thrown to the ground by Secret Service agents. In this erudite and learned collection of essays on the meaning of authority, Bruce Lincoln, a professor of religion, compares this challenge to authority with others like it in classical antiquity and Norse medieval history. Rather than dwelling on the "crisis of authority" (from the perspective of the ruler), or the threatening expansion of state authority (from the perspective of the ruled), Lincoln focuses on the relations between authority figures and their subjects. How do they communicate with one another? What is the evidence of the legitimacy? He focuses on a variety of situations in which authority figures are heckled, confronted, or challenged: Agamemnon is heckled by Thersites, Caesar's authority was to be challenged by a prophet, and, from a 10th-century case, the Norseman Egil mocks the Norwegian king Eirik.

What makes these events so disruptive is that they occur within privileged places: the Greek agora, the Roman Senate, the medieval Norwegian Thing. He makes an interesting point that while it is easy to argue that violence puts an end to opportunities for discussion, physical force should not be objectified and privileged in this way: threats are actually an implicit part of much discourse, and focusing only on the situations in which these threats are realized denies the importance of the threatening situations. Violence cannot be understood as something that only affects those whose bodies feel the pain: it affects and terrorizes all those who hear about it third hand.

There are a number of problems with the book. One is the relatively uncritical use of source material. While the textual sources of Lincoln's information are an epic poem, a Scandinavian saga, and the works of Roman historical writers, no mention is made of the way in which the genres themselves might influence the ways in which authority is presented.

Another shortcoming is the relative lack of attention to the social, historical, and comparative context of the different texts. There is little discussion in these pages about the historical differences among these cases or about the fact that the differences might have something to do with the contexts—historical and sociocultural—in which the performances take place. Nor is there any exploration of counterexamples to the ones presented here. Why these cases? How are they related to each other? How can we explain the differences among them? Why did authority develop

in a different way in Norway than it did in Greece or Rome? Instead, readers are provided with a virtuoso reading of some sample texts, and some stirring moral and analytical conclusions are drawn.

Finally, Lincoln's scholarly engagement with other theorists who write about authority is limited. While he suggests that he has been influenced by Bakhtin and Foucault, he does not explain the ways in which these two—very different—writers have separately influenced him and their relation to his work. In fact, the only scholar whose work he critically engages is that of Hannah Arendt.

Nevertheless, the book is engaging, particularly when it is focused on individual texts. It will be of interest primarily to humanities-oriented scholars of authority.

The Ethics of Catholicism and the Consecration of the Intellectual. By André J. Bélanger. Montreal: McGill-Queen's University Press, 1997. Pp. viii+242. \$44.95.

Lewis A. Coser
Boston University

André J. Bélanger, who teaches political science at the University of Montreal, here embarks on a significant intellectual venture. But, alas, the results of his labors are highly unsatisfactory. He wishes to show that Catholic countries, mainly France, have developed a highly influential circle of intellectuals by secularizing the Catholic medieval tradition of Thomas Aquinas and other church fathers, while Protestant countries, mainly Great Britain, lacking a comparable tradition, could never develop an intellectual grouping similar to their counterparts on the other side of the Channel.

The trouble starts with the author's definition of intellectuals as individuals who, "as producers of signs and symbols transcend her or his particular or professional sphere of competence and become actively involved in the political arena" (p. 83). There have been many debates on the definition of intellectuals, but to my mind nobody has heretofore made political involvement a requisite for being an intellectual. Are Sigmund Freud or Friedrich Nietzsche not to be counted as intellectual giants of the modern world because they avoided political engagements? Alas, the author's definition misses a high proportion of contemporary or past intellectual leaders as well as followers.

But the trouble goes way beyond issues of definition. Bélanger argues that, in countries with a strong religious tradition and in which the Catholic clergy long dominated intellectual life, the eclipse of clerical spiritual guidance in modern days led to the emergence of lay strata of intellectuals of a secular cast of mind. In Protestant countries, on the other hand, stress on the "priesthood of all believers" allowed no such clerical domination to prevail. As a result, increasing secularization did not produce indepen-

dent secular circles of intellectuals. There may be some superficial insight here until one thinks of the vivid intellectual life of Goethe and his Weimar circle—but then Goethe was not very much interested in politics. And when it comes to the modern Weimar Republic, one encounters as many major intellectual figures in the Protestant North as in the predominantly Catholic South. Bélanger is fascinated by Jean Paul Sartre and his circle in post-World War II Paris. And I will admit that traces of Jesuit thinking styles can be discerned in the writing of the French existentialists. But what of the worldwide impact of Ibsen, Kierkegaard, and Strindberg from the Protestant rim of Europe?

My trouble is not only with the above substantial matters but also with the author's mode of reasoning. The book consists in the main of brief sketches of the moral philosophy of men and women authors. Some are given just a few lines, fewer are given a paragraph, and still very few others rate a page or more. The treatment of these figures is on the level of elementary introductory textbooks. What earthly use can there be for a synopsis of the Abbé de Reynals seven-volume history of the contact between Europe and the Indies in one paragraph? Can Coleridge's profound thoughts about a modern clerisy really be done justice on a little over a page of text?

But there is more: Bélanger provides us with short summaries of the work of many thinkers, but he omits any reference to the times, the cultural configurations, and the historical locations of the authors he summarizes. The treatment of these matters on Bélanger's page makes it appear as if Karl Mannheim had never existed, and as if the *Annales* school of French historiography had never gotten off the ground. How can we afford, standing on the shoulders of such giants, not to celebrate their work by following in their footsteps? How can we afford to write about Thomas Aquinas or any of the major medieval churchmen without consulting the work of the *Annales* historians?

To summarize: (1) the author's contention that to be an intellectual entails political engagement fails miserably when confronted with the evidence. (2) His thesis that intellectuals in the Catholic world have an advantage over their Protestant colleagues by virtue of their developing their thought in the line of descent of their clerical forefathers fails any empirical proof. (3) An attempt to write intellectual history without having recourse to the tradition of the sociology of knowledge can only be seriously deficient.

Essays on Religion. By Georg Simmel. Edited and translated by Horst Jürgen Helle in collaboration with Ludwig Nieder. Foreword by Phillip E. Hammond. New Haven, Conn.: Yale University Press, 1997. Pp. xx+223. \$27.50.

Lawrence A. Scaff
Pennsylvania State University

With the welcome publication of this collection of essays, sociologists and students of religion now have available in a single volume everything significant that Georg Simmel wrote about religion. The 11 essays Horst Helle and his collaborator, Ludwig Nieder, have translated cover the two decades of Simmel's mature writings, from "A Contribution to the Sociology of Religion" of 1898 to the concluding "Conflict of Modern Culture" of 1918. Most important, the translators also include a new complete translation, covering nearly 80 pages, of Simmel's most sustained discussion of religion: *Die Religion* (1906; 2d ed. 1912), originally published in Martin Buber's popular multivolume series *Die Gesellschaft*, which was dedicated to discussing different aspects of modern society. (Tönnies's study of *Custom* and Sombart's book *The Proletariat* were also written for this series, which Buber called "a collection of sociopsychological monographs.") The lengthy reflection on *Die Religion* is placed at the end of the translation, thus standing as a summation of Simmel's views, while the 10 shorter essays are conveniently grouped under four headings that give definition to Simmel's interests: religion and modernity, personality, art, and methods for studying religion.

The subject matter in these writings is not simply the sociology of religion as a field for observing "interaction," but more particularly the "lived experience" (*Erlebnis*) of the religious person. As Phillip Hammond points out, the essays investigate the sphere of religiosity or "religiousness," and thus in contrast to Max Weber's work reveal the "religiously musical" side of Simmel's perceptions. Following others, such as William James in *The Varieties of Religious Experience*, Simmel distinguishes between religion as a doctrine or "set of claims," and religion as a "form of life" or "attitude of the soul," selecting the latter (Kant's phrase) as his primary concern. But, unlike James, Simmel explicitly repudiates the "mechanistic psychology" that views religiosity as a mental "entity," instead proposing to see it as "something inward and personal" and a "form" through which we experience life (pp. 5, 104). Religiosity is thus a type of affective relationship, an aspect of the dialectical struggle between "life" and "form" that Simmel places at the center of his overarching theory of culture. And it is religiosity that provides the essential dynamic of religion as a set of claims. In Simmel's words, "Just as cognition does not create causality, but instead causality creates cognition, so religion does not create religiosity, but religiosity creates religion" (p. 150). Or as he expresses the idea in another passage, the logical sequence is from religiosity to "social phenomena" to "objective religion" (p. 211).

Reduced to essentials, the problem Simmel then faces in these pages is twofold: What can be the meaning and experience of religiosity in the modern age, where we are faced with a withdrawal of the "content" of religion, while the "religious need"—that is, the form called religiousness—remains with us? How are we to understand the reality of a religiosity that exists *independent* of any content, that has been transformed from "a transitive to an intransitive activity" (p. 23)? Second, what is the consequential connection between religiosity and social phenomena? Can there be instructive parallels or analogies between religious conduct and social interaction? Simmel's complex responses take different forms themselves. But in general his reflections on art, subjectivity, inwardness, or "reconciliation" within the personality trace the engagement of the self—especially the contested modern self—with religiosity. In addition, especially in the longest text, *Die Religion*, his thesis is that religiosity is itself a *social* form, observable within other nonreligious relationships, serving as a template for right conduct, triggering the search for social unity in a "higher order," or modeling the relationship of the individual to the group. These passages show Simmel at his best, revealing from his unique perspective the problems a sociology of religion should pursue.

Helle and Nieder have produced an elegant and exceptionally readable translation of Simmel's texts. The paradoxes and powerful conceptualizations in Simmel's original are cast in a form that will give the careful reader access to the complexities of his thinking. Having said this, it is probably still the case that Simmel's German original possesses greater tension and conceptual depth than any translation can fully match. For example, to render both *Dasein* and *Existenz* as "existence" in the opening passages of *Die Religion* (p. 138) misses a distinction between fundamental being and transient existence that Simmel's language points toward, thus anticipating one important line of thought in 20th-century philosophy. But this is an issue no translation can fully resolve. The editors have wisely chosen to promote a clarity that does justice to Simmel's overall intentions, meaning, and argument, and that provides students of modern society and religion with an accessible body of thought by one of sociology's key founders.

The Order of Rituals: The Interpretation of Everyday Life. By Hans-George Soeffner. Translated by Mara Luckmann. New Brunswick, N.J.: Transaction Publishers, 1997. Pp. xvii+176.

R. S. Perinbanayagam
Hunter College, City University of New York

In this work Soeffner addresses two fundamental issues. To begin with, he dismisses certain recent claims about the "death of the subject" and affirms that his purpose is to "pursue selected paths to the contemporary

belief or hope in an autonomous human subject" (p. II). This subject, he argues, can be found in the interactional rituals by which individuals constitute themselves and allow themselves to be constituted by others. Ritual for Soeffner, in a way reminiscent of Goffman's use of the term, are "symbolic forms of (self) presentation" (p. viii). To the extent that an individual uses these rituals to engage in "self observation of social presentation and self staging," the individual realizes himself. To the extent that we share and constitute these rituals into lifestyles, they "signal not only consumer habits but also membership in collective attitudes about conduct and value systems" (p. vi). This constitution and presentation of self in and through rituals also answers, for Soeffner, the problem of macro-order insofar as "the community forms and passes on the rituals, thereby stabilizing itself" (p. 74). Further, the use of "interaction rituals establishes the range of possibilities for creating temporary interactional communities where all participants maintain a pretty much implicit and anonymized context of order for social action" (p. 74).

Soeffner seeks to support these claims by examining a variety of cases beginning with a study of the rituals by which Martin Luther constituted himself and his "subjectivity." Luther's transformation of a collectively oriented faith to an individually oriented Protestantism is examined in terms of the everyday rituals and discourse by which it was accomplished. Soeffner traces Luther's path with material from a variety of sources, and I will give here the analysis he provides of the transformation that Luther achieved of the sacraments of the Roman church. He argues that the sacraments function to achieve an "organizing and forming of the individual Christian life into stations." This organization is guaranteed by the community of the faithful whose agent is the "religious specialist" and the parents and godparents who participate in the christening and confirmation. "The institutional repetition of the rite in Christian mythology thus turns into . . . the faith in the magical character of repetition: the repetition of myth becomes the myth of repetition" (p. 13). These "ritual network of rules and regulations," Soeffner argues, and the "protective distance that the network lays between the faithful and the immediacy of the sanctified" are what Luther sought to destroy. For him the immediacy of the "manifest word" and the immediacy of "God himself" was everything. Luther pushes aside the distancing features of the Roman practices and institutes methods by which "genuine" access to the "true word" and to the immediacy of "God's given word" can be made. One of the ways he achieves this is to reduce the sacraments to those concerned with baptism and the Eucharist. This move of abolishing one set of rites only to replace it with another made it possible for believers to construct a Protestant, different self (p. 13). As Soeffner puts it, "To the extent that Luther's teachings are lived, and his instructions for leading a Christian life are perceived and implemented in everyday life, a new type of lifestyle emerges . . . as does a new type of sociality: isolation rather than community" (p. 21).

In examining a contemporary process, the "punks" achieve style and

stylization by a process of "the superelevation of everyday life." In Soeffner's words, "A specific historical style, is first of all an observable (self-) presentation of persons, groups or societies. Style as specific presentation marks and manifests the individual's belonging not only to a group as community, but also to a specific demeanor and form of life which these groups or communities felt compelled to maintain" (p. 50). The punks as a group achieved this stylization by adopting distinctive hairstyles and speech patterns and by undertaking various activities in public that amounted to a ritualized parody of those of the parent society.

This argument is carried forward with a study of certain political demonstrations presented by religious groups, feminists, trade unionists, homosexuals, "gray panthers," and so on, who did not have a durable and centrally run parent organization. Nevertheless, the members presented themselves in distinctive clothing, wearing emblems, carrying posters, and performing various ritualized gestures. These movements initially set out to defy established customs and traditions and to reject rituals and ritualism, but they ended by inventing their own rituals. The conclusion is clear: it is impossible for symbolic creatures to avoid the ritualization of everyday life.

Drawing more explicitly from Durkhiem, Soeffner next tackles the rituals practiced by the pigeon-breeding miners of the Ruhr District in Germany. Leading a molelike existence underground in the mines they emerge from them to look into the sky and fly pigeons and feel a sense of liberation. Breeding, keeping, flying, and discussing pigeons create a sense of community among the miners. The pigeon becomes a quasi-totemic entity, a "common symbol" that allows a number of people to engage in various actions and to share a sense of common purpose. In other words, the activities connected with the pigeons undertaken by the miners are rituals of self-presentation that establish an order of relations with others.

In the last two chapters Soeffner examines the moves made by the Dortmund beer company to present their product to a mass audience in the best possible light on television by parodying the voting system of parliamentary elections and the interactional emergence of populist and charismatic leaders in the ceremonies held at Verdun by German and French war veterans. In every way this is a fruitful examination of the rituals of everyday life and addresses, with imagination, evidence, and sound argumentation, some of the perennial questions of sociology.

The Soul of Development: Biblical Christianity and Economic Transformation in Guatemala. By Amy Sherman. New York: Oxford University Press, 1997. Pp. xviii+214. \$45.00.

David A. Smilde
University of Chicago

In *The Soul of Development*, Amy Sherman argues that rapidly expanding Evangelical Protestantism in Guatemala fomented a variant of the "Protestant ethic," which is already having a positive effect on socioeconomic development. As unfashionable as it may sound, this is an entirely reasonable thesis and worthy of research. To pursue it empirically, Sherman interviewed a stratified random sample of individuals from various tendencies within Evangelicalism and Catholicism. The product is a series of statistically significant correlations between adherence to conservative Christian doctrine and rejection of traditional animistic beliefs, receptivity to entrepreneurialism, educational attainment, acceptance of meritocracy, and acceptance of literacy for both men and women.

Despite acknowledging that her correlations "cannot prove cause-and-effect relationships," Sherman argues that "in light of the consistent weight of evidence uncovered through my fieldwork . . . and various ethnographic studies, I have chosen to employ causal language" (p. 74). This move is something like fantasizing that you are eight feet tall instead of practicing your free throws but is especially unacceptable here for the following reasons. First, Sherman presents her quantitative research as a remedy to the anecdotal (a term she freely interchanges with "qualitative") findings of ethnographic studies (pp. 12, 50). Second, when one looks at the studies she mentions in chapter 2, almost without exception, they speak in terms of correlation and not cause. Third, Sherman passes over the numerous studies that show little or no effect on socioeconomic well-being, while she is completely uncritical of studies that support her position (a pattern repeated throughout the book). For example, she cites what she calls Elizabeth Brusco's "finding," that a male's becoming Evangelical permits 20%–40% of household income formerly spent on alcohol to be redirected to family needs (p. 46). When we turn to Brusco's book, we see that this "finding" is actually a reference to an estimate Orlando Fals Borda made over 35 years ago in his study of one Colombian mountain town (Elizabeth Brusco, *The Reformation of Machismo* [University of Texas Press, 1995], p. 176, n.4). Fourth, Sherman's "fieldwork" does not get to questions of causality either. The only data that address "before and after" issues are individual converts' own testimonies of how their lives have changed. These are interesting in themselves but poor sources of biographical facts.

Curiously, multiple regression could have easily been used to analyze Sherman's data and *approach* causal statements. At least by controlling for intervening variables, one can eliminate *some* alternative hypotheses. She in fact mentions that she tried this with income, ethnicity, occupa-

tion, and years of education, and that "frequently," the statistical significance remained (p. 74). Nevertheless, this data analysis is neither presented nor ever mentioned again. The reader can only wonder why.

Indeed, if taking correlation for causation were a plausible move, there would be little debate on Protestantism and social change. For the better part of the 20th century, arguments portraying Protestantism as an important factor in the development of the capitalist economy and democratic polity have been countered by arguments that Protestantism is an epiphenomenal, after-the-fact rationalization of the location of individuals and groups in structural transformations already underway. What few scholars contest, however, is the high correlation between Protestantism on the one hand and capitalism and democracy on the other.

Were Sherman's thesis supported by data actually relevant to it, the argument it is placed in is still flawed in several ways by misuse of the term "development." Despite noting the ambiguity of the term toward the end of the book, she uses it in an unreflective way throughout. All of her analysis is at the microlevel, yet she continually speaks of Evangelicalism fomenting "development" at the national level. Even if one were to agree with Sherman's view of the microlevel effects of Evangelical Protestantism, one could have a negative assessment of its long-term consequences if starting from a different theory of macrosocial change—for example, one could reasonably argue that individualistic gain works to demobilize the social movements that might produce the structural changes, at the local, national, or international level, necessary to overcome Latin American social problems (see, e.g., Steve Brouwer, Paul Gifford, and Susan Rose, *Exporting the American Gospel* [Routledge, 1996]). Furthermore, her analysis of the would-be significance of the data is clearly erroneous. She maintains that "research that provides evidence of the influence of cultural factors on development prospects is valuable for reinforcing the emerging shift toward neoliberalism in development thinking" (p. 17). Were we to accept the claims she makes for her data, their relevancy to "development thinking" would be quite the opposite. If Evangelical Protestantism supports development, while the "Cristo-Pagan worldview" (her term for folk Catholicism) hinders it—and the great majority of Latin Americans are now "Cristo-Pagans" and will remain so in the foreseeable future—then, by the force of her own argument, neoliberal development policies will be self-defeating.

In sum, while the analytic intention of this book is entirely reasonable, it is incompetently researched, chauvinistically presented, and erroneously analyzed. Virtually the only scholarly use I can imagine for it is as a datum for those studying the construction of "otherness" in North-South relations.

Planting Trees in the Developing World: A Sociology of International Organizations. By Steven R. Brechin. Baltimore: Johns Hopkins University Press, 1997. Pp. xv+254. \$48.50.

Frank Hirtz
University of California, Davis

Organizational analyses of agencies of development assistance are rare indeed. For that reason alone Brechin's study deserves attention. The book deals with the way in which one specific type of development assistance is put into practice by three distinctly different organizations. Brechin depicts differences in organizational setup, management style, and the implementation of community forestry programs in the World Bank, the Food and Agricultural Organization of the United Nations (FAO), and CARE USA, and develops indicators for measuring such differences. The key indicators are technology (skills possessed and procedures used by an organization), organizational environment (influential forces outside the organization proper), and structure (the organization's division of labor). These indicators allow Brechin to ascertain the reasons for the markedly different performances of these organizations in community forestry programs. Based on this empirical foundation, the book's two main ambitions can be spelled out—first, to give a detailed account of the connection between the organizational structure of the agencies and the way in which these agencies perform and, second, to contribute to the theoretical body of sociological literature on organizations, especially by attempting to create a causal linkage between organizational character and performance. Thus, the project has the format of a quasi experiment, in which policy implementation is kept constant and the execution is made variable.

The book is meticulously organized. A data-packed introduction develops the appropriateness of the key indicators and provides the background against which the next three chapters can unfold their richly detailed accounts of each of the organizations' general structuring philosophies, modes of operation, and political embeddedness. Anyone who is not familiar with the maze of international organizations in the field of development will appreciate the balanced and very knowledgeable overview. Two further chapters then ascertain how each organization stacks up against the three key criteria developed in the introduction: commitment to rural development forestry, appropriateness of the work, and general effectiveness. In a final chapter, Brechin moves "Toward a Sociology of International Organizations."

Yet, with all this clarity and an impeccable knowledge of community forestry and international organizations, the book is only partially convincing. On the positive side, Brechin is very familiar with the theoretical literature on organizations and deftly juxtaposes this vast amount of literature with the three case studies at hand. Thus, in part this book reads as a plea to development practitioners (and others as well) to learn that

it is worthwhile to think theoretically about their applied work. Conversely, it often sounds as if Brechin admonishes the organizational theorists who dare to venture into the underbrush of international organizations. Yet, at the end I ask myself whether the original question (How is it that different organizations do the same thing so differently and what do we make of that difference) is so promising after all. I will use a metaphor to explain my puzzlement: It is as if one were to ask why different instruments sound different and do not play together, even though all instruments are used to play music. Is it not a moot question to ask members of a string quartet why they are not playing in a big orchestra? It seems equally moot to ponder why a trumpet does in fact sound different from an electric guitar. In thinking about these questions, I learn a lot along the way about the individual instruments, various musical styles, and the fascinating possibilities of orchestration. Yet, in the end a trumpet is still a trumpet and the drums are still louder than a harp. One should not stretch metaphors too far, but I hope it helps to make the following point clearer: organizations that deal with a policy output (as opposed to a commodity to be consumed or a service to be rendered) are themselves a reflection of the prevailing political ambience; they occupy a place defined by a host of political factors all of which delineate the arena of activity and the level of competence commensurate with that very space they occupy. Obviously, within these arenas of activities and competence, important variations take place. That Brechin has shown us these variations is his prominent contribution. That he cannot show that this variance is not explained through the details of the different organizations diminishes the strength of his theoretical framework.

I presume the problem for Brechin has its origin in his ambivalence about distinguishing clearly between an evaluative study and presuppositions from an analytically and theoretically driven question. In writing an evaluation and telling the organizations being studied what they are doing and what they in fact can do better, even within their accepted parameters, Brechin clearly puts them under a sharply focused magnifying glass. All three organizations can learn a great deal from his analysis, be it their limits or their untapped potentials. It is one thing (and a good one at that) to let an evaluation be guided by organizational theories, but it is different from claiming that an organizational evaluation is a theoretical contribution.

Narrating the Organization: Dramas of Institutional Identity. By Barbara Czarniawska. Chicago: University of Chicago Press, 1997. Pp. vii+234. \$45.00 (cloth); \$15.95 (paper).

John W. Mohr
University of California, Santa Barbara

Institutionalists have shown us that organizations are suspended in webs of meaning. Myths, symbols, and the struggle over what will count as legitimate are the very stuff out of which organizational phenomena—transactions, strategies, resources, and power—are constituted. Yet institutionalists have tended to shy away from a direct confrontation with meaning. Being scientists, not humanists, their goal has been to show how meanings operate, not to decipher their content. This is a temporary solution at best. To fulfill its promise, institutional theory must find a way to interpret the meanings that make up organizational life and to funnel those interpretations into the machinery of social science. Barbara Czarniawska's new book provides a promising beginning for such an endeavor.

Czarniawska proposes that we retheorize organizational processes as social narratives—stories that are told about what organizations are, how they operate, what they are attempting to accomplish. She argues that meaning is constructed and conveyed in narrative form. We carry stories around in our heads that explain the workings of the world to ourselves. This enables us, in a given context, to answer the question, What is going on here? For over a decade now, cognitive scientists have recognized the importance of narrative devices as mechanisms for structuring the organization of information and they have dedicated much research to analyzing these processes. While acknowledging the importance of this work, Czarniawska relies more heavily on Kenneth Burke's dramaturgical theories of action and, especially, on the insights of literary theorists, who, after all, have been in the business of understanding the nature of stories for a very long time now.

For social scientists, there are many advantages to this approach. First, thinking narratively can enable us to reduce the complexity of interpretive work. Narratives clump meanings together. There may be a great many different meanings in the air, but relatively few stories to gather. Second, narratives provide a basis for analysis. Meanings fit into story structures. They fulfill narrative functions by advancing the plot or echoing a theme. Understanding the storied character of organizational existence thus provides us with a means of studying meanings in a systematic, rigorous, and structurally organized fashion. Third, the shift to narrative thinking opens up many alternative avenues for theorizing organizational behavior. Problems of structure and agency, conformity and innovation, strategy and conflict appear quite different when seen from within the framework of narrative theory than when they are portrayed according to the logic of physically or biologically oriented metaphors (fields, niches, etc.).

The book is divided into three parts. The first section provides an introduction to the theory and logic of narrative analysis. Here, Czarniawska provides a crisp and fascinating summary of the broad literature on narrative theory as well as an intriguing account of previous work on narrative analysis in organizations.

The second section of the book contains examples. Drawing upon her own fieldwork among Swedish social insurance agencies, Czarniawska shows how organizational actors understand their own activities and the activity of those around them in terms of overarching narrative constructs. Her examples include various attempts to impose a new inter-organizational budget and accounting system, to reform the tax structure, to privatize public agencies, to shift the burden of financial obligations from the public to the private sector, and to bring a new computer system on line. Some of these events unfold according to relatively bounded storylines. Others proceed under the aegis of sagas, or what Czarniawska refers to as serials—thematically related episodic structures that may never reach closure. In either case, recognizable genre conventions enable participants to pull together around a common set of interpretive expectations.

In the third section of the book Czarniawska critiques dominant organizational theories, especially the institutionalists, and shows how narrative analysis can lead to conceptual improvements. Some of her suggestions include a proposal for analyzing the kinds of genre conventions that obtain within a given organizational field, a project that might allow us to see how certain types of narratives are promoted while others are disallowed. She notes how literary conventions for analyzing paradox can be applied to explaining the role of organizational ambiguity and information uncertainty. And, she highlights the work which has been done by anthropologists on multivocality in narratives, proposing some intriguing parallels to problems of power in organizational theory.

Whether narrative analysis will prove to be the next stage of development in institutional theory remains to be seen. Nonetheless, Czarniawska's book is certainly an eye-opener. It is filled with creative rethinkings of traditional problems of organizational theory and with provocative suggestions for research. It should be of interest to all forward-thinking organizational researchers as well as any scholar interested in the problems of narrative analysis.

Rationalising Medical Work: Decision-Support Techniques and Medical Practices. By Marc Berg. Cambridge, Mass.: MIT Press, 1997. Pp. x+238. \$30.00.

Stefan Timmermans
Brandeis University

An emergency department nurse enters patient information in a computer system. The computer screen directs the nurse what questions to ask and what vital signs to record. At the end of this brief interview, the computer flashes a diagnosis on the screen. According to the computer system, the patient's chest pain is likely an indication of serious cardiac problems, and the patient needs to be admitted immediately to the intensive care unit. Depending on whether one is a critic or an advocate of computer decision-support systems, this scenario is another instance of either medical nemesis or the next revolution to rationalize medicine. Advocates of medical information systems argue that the presumed objectivity of rational tools will improve an imperfect medical practice, while critics claim that attempts at creating and installing decision-support systems are doomed because they can never capture the infinitesimal complexity of human practice and that every attempt at decision-support systems inevitably reduces the richness of human skills and interactions.

In his thought-provoking study, *Rationalising Medical Work*, Marc Berg argues that both sides of the debate miss the point because they separate tools from practices. According to Berg, rationality is an achieved quality, the outcome of a lengthy process of converging heterogeneous needs and negotiations. Berg analyzes how three forms of rationalizing medical practice—drug protocols, computer-based decision-support tools, and clinical decision analysis—have influenced medical practice. He first looks at how the image of a "rational" medical practice emerged in editorials of leading medical journals concurrently with the development of specific tools since the Second World War. The solution of rationalizing tools was constructed simultaneously with the problem of an inadequate medical practice.

Turning to the toolmakers themselves, Berg analyzes next how the designers of the rationalizing tools obtained a smooth fit between their tools and practice by redescribing medical practice in the image of the tools. Toolmakers try to discipline medical practice with guidelines, prerequisites, and a reshuffling of existing professional relationships to make their devices work. Finally, Berg analyzes the ad hoc work that needs to be done to keep the rationalizing tool on track in the hospital. Berg's image of medical practice is endlessly fluid and open-ended. Tools work but not the way their designers intended because the tool's purpose is just one element in the management of patients' trajectories and needs to be weighted with the tasks at hand.

Protocols are ubiquitous in medicine and work where they should not be according to the critics, and computer decision-support systems are

shelved away when they should have revolutionized medicine according to their advocates. Critics and advocates miss the point because they solidify tools and practice. According to Berg, tools do not just enter or modify a preexisting practice, medical practice and tools shape each other. They emerge simultaneously to constitute new configurations, with new forms of knowledge, new challenges to keep both the tools and the users on track, and ultimately new opportunities for diagnosing and caring.

Berg convincingly argues against conventional wisdom. His empirical material is rich, diverse, and rigorously analyzed. He impressively draws from ethnomethodological studies of science and the French actor-network theory to make his points. His study is located at the intersection of sociology of health and illness and science and technology studies. Sociologists of health and illness have regularly invoked the proliferation of "advanced technologies" as the residual explanatory variable—the magical *deus ex machina*—to explain all kinds of changes in medicine (from the rise of health care costs to the alienation of dying in hospitals) during the past 50 years. At the same time, the same sociologists rarely bothered figuring out how these new tools exactly impacted the work of doing medicine. Berg challenges sociologists of health and illness not to employ technology as an isolated factor but to study it as an intrinsic part of medical practice. In addition, his study provides science and technology scholars with provocative insights in the difficult transition from the locale in which new tools are designed to the places where they are used. Berg is one of the first to take Bruno Latour at his word and follow decision-support systems in action; he starts with the idea of rationalizing medical practice and ends with the mutual shaping of tools and patient care in hospitals.

Berg's outstanding analysis remains largely on a micro level. Although this lies beyond the scope of the book, I wondered how these new tool-practice configurations resonate on a macro level. In times of health care turmoil with managed care experiments, exploding health care costs, concerns about health care quality, and legions of the uninsured, how do computer-supported medical decision systems and protocols tip the political power balance of health care providing? In the end, who benefited and who is disempowered with the rationalization of medicine? The answers to these questions are probably less straightforward than conventional wisdom would suggest, and we can only hope that Marc Berg will tackle them in turn.

Bodies in Protest: Environmental Illness and the Struggle over Medical Knowledge. By Steve Kroll-Smith and H. Hugh Floyd. New York: New York University Press, 1997. Pp. 217. \$24.95.

Judith Green

London School of Hygiene and Tropical Medicine

In *Bodies in Protest*, Steve Kroll-Smith and H. Hugh Floyd tell the stories of people with multiple chemical sensitivity (MCS), people struggling to live in a modern world where everyday material culture (perfumes, ink, computers, televisions, synthetic clothing) literally makes them sick. The reality of MCS is contested. While the majority of clinical medical practitioners deny its legitimacy as a diagnosis, sufferers face a daily challenge either to escape their hazardous environments or to persuade others to make them safer. If these stories were merely of conflicts between the life world of lay accounts and the biomedical expert system that cannot heal, there would be little new here to add to existing literature on living with chronic illness. However, Kroll-Smith and Floyd argue that the unusual, and sociologically interesting, ramifications of MCS are that it disrupts what is known about both environments and bodies; namely that environments are either safe or unsafe (extreme) and that the physical body can be taken for granted. With MCS, the most mundane of environments becomes dangerous, and bodies in those environments become unpredictable and disordered. Biomedicine constructs and acts upon universal and predictable bodies, but the bodies of people with MCS react in individual ways that can only be discovered through personal experience itself. Kroll-Smith and Floyd somewhat contentiously claim that bodies thus become "sources of unmediated knowledge" (p. 93). This experiential knowledge is combined with the language of biomedicine to make a coherent and convincing story for others as well as the ill self. Sufferers do not reject expertise, rather its practitioners, as they utilize a technical vocabulary to forge a practical epistemology that is both rational and instrumental. MCS is interesting as an exemplar of some key concerns of late-20th-century Western society: the environment, the democratization of expertise, and the "rediscovery" of the narrative as a legitimate mode of knowing. In addition, these accounts of MCS resonate with contemporary sociological themes—such as the body and the nature of risk and uncertainty in late modernity—and there are hints throughout the text that a close attention to the narratives of sufferers will lead to fundamental insights for sociological theory.

The anthropological tension between narration and analysis, between "telling it like it is" and deconstruction, is here solved through realism and a belief that (contra-Foucault) bodies do preexist culture (p. 197) and that we can experience them directly even if the classification of those experiences is cultural. Thus bodies undercut theory. Our expectations of what the implications of this might be are raised by an initial claim that "MCS is a critique of both modernity and postmodernity" (p. 12),

but we are already inundated with critiques of modernity and the authors' critique of postmodernity remains undeveloped. Instead, we are presented with a "success story" of the power of lay people to mobilize within a pluralist democracy, evidenced by instances of workplace adaptations, "fragrance free" seating in public places, and the recognition of the reality of MCS within popular culture.

Kroll-Smith and Floyd fundamentally remain advocates for those whose stories they recount rather than analysts of them. Their aim is "to make the 'other' . . . familiar" (p. 13), and they are pleased to report that none of those interviewed for the study disagreed with their retelling. The "other" is indeed made familiar, even reassuring, in the book's liberal conclusions—that in achieving social change, the most effective strategy is to adopt the rhetoric of the "center" (biomedical discourse and a legal discourse of rights and citizenship, utilizing the gains of the disability movement in the United States) and to move incrementally toward a more inclusive society.

A less realist approach would have opened other, if less comforting, lines of analysis. Why is it possible, now, to speak of environmental illness and for us to experience our bodies as porous, as individual, as separate from ourselves? Kroll-Smith and Floyd reiterate their key point that MCS has upset our assumptions that environments are either safe or unsafe, but this point has already been made (to a sociological audience at least) by the writings of Ulrich Beck and others, in claims that uncertainty is a key feature of late modernity (in which we all face a multitude of risks), and where science is as likely to pose dangers as it is to provide solutions. Equally, there is little new in a questioning of the distinctions between "lay" and expert discourse. Accounts of illness are in rational and medicalized terms because this is what is possible to say and to hear. An equally plausible story would have been of MCS as having been made possible by the uncertainties of late modernity, rather than being disruptive of it. Neither the existence of MCS nor this analysis moves theory forward as much as was promised, but it is perhaps churlish to expect that as well as a detailed narrative account. What Kroll-Smith and Floyd have done, with both clarity and sensitivity, is provided considerable insight into an important arena of contemporary experience and rich data for the postmodern theorists to mull over.

Community Policing, Chicago Style. By Wesley G. Skogan and Susan M. Harnett. New York: Oxford University Press, 1997. Pp. x+258. \$29.95.

Peter K. Manning
Michigan State University

Chicago has produced a pride of sociologies. It remains the home of urban sociology and applied social research with an edge. Skogan and Harnett offer here yet another exemplary, extraordinary, study that incorporates

both fascinating particulars and generalizable findings. It is well designed and subtly orchestrated and should become required reading in the social sciences. Not only is it engaging and clearly written, it contains hopeful notes about the future of urban policing.

The research question concerns the effectiveness of a community policing program in Chicago. "Community policing" is an elastic notion stretched to include everything from a social movement transforming policing to a highly circumscribed and delineated tactic. The authors call it a strategy, or principled means of allocating resources to achieve the ends of policing, that leaves the emergent tactics to practitioners. They identify four principles of community policing (pp. 6-9): policing that is (1) decentralized, (2) problem oriented, (3) responsive to citizens with respect to tactics employed and priorities set, and (4) committed to joint solutions to neighborhood crime and order problems. Community policing is a reaction to the radio-dispatched distant and mobile patrol strategies associated with the "professional" model of policing developed by O. W. Wilson, August Vollmer, and others. Through a variety of tactics, it intends to reorganize policing to reduce social distance between citizens and the police.

Unlike these, many reformist authors have promoted community policing, claiming (among other things) that it reduces costs, violence, calls for service, and crime levels even while it increases efficiency, satisfaction with the police, and police morale. Skogan and Harnett properly consider these as empirical questions.

The authors note the obstacles to change in policing, but unlike many books that purport to promote reform, they do not call for change in the occupational culture, unionism per se, the inspectorial organization, nor dispensing with middle management (e.g., the rank of captain was abandoned in Chicago). Given these constraints, the modest success of the Chicago program issues directly from the sponsorship and support of the mayor's office, newly budgeted resources, and the recommendations of a consulting group that advocated community policing in conjunction with the Chicago department's reorganization plans.

Chicago announced a community policing program (called "CAPS") in the fall of 1992; it was initiated in five experimental districts in April 1993. Skogan and Harnett, members of a multiuniversity team evaluating this project, summarize here the findings through the end of 1994. The project was funded by foundations and federal and state money. They interviewed 7,000 officers (in two waves) and a panel of 2,579 citizens at two points via telephone. In addition teams observed community beat and advisory meetings in the five prototype districts and studied and gathered observational material. The evaluation team continues to monitor the program, now into its fourth year and ostensibly citywide.

The CAPS program is unique. Although it was implemented in a dynamic, racially divided city, characterized by the authors in a succinct chapter entitled "Police and Politics in Chicago," CAPS clearly addresses problems generic to urban America. Community policing "Chicago style"

aimed to reduce specialization, to fix beat assignments, to increase and improve training to develop performance and evaluation measures, and to introduce crime analysis and crime-mapping facilities and integrated police-driven city services linked directly to the mayor's office. Concern for efficiency (business professionals supervise budgeting and management in the 25 police districts) and reorganization remains. On balance, Skogan and Harnett believe the program had an impact, although the crime analysis and mapping, integrating specialized units into the CAPS program, and performance evaluation were never implemented fully.

Surveys of officers revealed that the police remain ambivalent; older, black and Hispanic officers are more supportive of community policing, while those in both the experimental and control districts felt that CAPS eroded their autonomy and reduced their satisfaction. They felt, on the other hand, that it increased their skills, their closeness to communities, and had a positive impact on the community. The detailed and fascinating material presented on beat meetings in chapter 5 reveals the abiding police wish to control, to take action, and their emphasis on patrolling and crime control as core functions. The meetings show clear evidence of differences in perspectives and styles, although problem solving did emerge. Physical disorder was reduced in one of the districts and crime in another. Chapter 6 is essential reading for any student of urban sociology or policing. It contains the first published examples of "coproduction" (a common buzz word in the community policing litany) of problem solving and reveals dramatically the messy world of volunteer action, and urban politics and organization.

The impact of the program on the five districts was remarkable—five of the 10 indexes of community impact were significant (Skogan's survey of six other Community policing programs across the country showed comparable results), and whites, Africans-Americans, and homeowners benefited most. Hispanics, in effect, did not receive the "treatment." They showed low knowledge, low participation, and low satisfaction. Although the authors present data comparing experimental and control groups, data on within-district differences could have been useful. Some of the most intriguing findings revealed differences in perceptions of police visibility, crime attack tactics, and excessive violence by race. For example, while the blacks surveyed liked visible police and felt the Chicago police were less violent as a result of CAPS, they reacted against crime attack and stop and search tactics.

Community Policing, Chicago Style masterfully sets out an experimental design, identifies the elements of the program, asserts a logical model by which program and results are connected, presents abundant data clearly, and is cautious about generalization and modest about findings. The CAPS program is well planned, well documented, and well implemented. It is closely connected to political resources and the mayor's office. Finally, it is about service, civility, and community organization, not about marginalizing the poor and unfortunate through slogans that jus-

tify urban terrorism (such as "zero tolerance"). One hopes for a series of such engaging and clear monographs from these authors.

Just Punishments: Federal Guidelines and Public Views Compared. By Peter H. Rossi and Richard A. Berk. New York: Aldine de Gruyter, 1997. Pp. xi+243. \$39.95 (cloth); \$19.95 (paper).

Jerry Van Hoy
Purdue University

Just Punishments reports the findings of a survey developed for the United States Sentencing Commission to compare public views on sentencing felons convicted in the federal courts with guidelines that were originally passed by Congress in 1987. The book contributes to a set of relatively small literatures on how the public ranks the seriousness of different types of crime and the public's sentencing preferences. The study is based on a factorial survey method in which weighted random samples of respondents were asked to sentence convicted felons described in short vignettes that randomly varied 20 crime types, 73 crime examples, and the characteristics of offenders. Rossi and Berk's main finding is that the general pattern of sentencing suggested by the public is similar to what is found in the federal sentencing guidelines. That is, despite variation in the specific magnitudes of sentences suggested by respondents and the guidelines, both tended to agree on the seriousness of criminal acts. For example, at the level of crime types both guidelines and respondents offered the longest sentences for kidnapping and drug trafficking and the shortest sentences for drug possession. The authors conclude that the evidence "support[s] the Durkheimian view that the laws of a society tend to reflect that society's normative consensus" (p. 210).

Unfortunately, there are several problems with the Rossi and Berk argument that sentencing guidelines reflect a public consensus. First, there are a number of areas of specific disagreement between guideline sentences and respondent sentences in their study. The most obvious example is drug trafficking. Though Congress has mandated highly punitive sentences for dealing crack cocaine (median guideline sentences of about 22 years), respondents treated crack as similar to dealing powder cocaine, heroin, and even marijuana (median sentences of 8–10 years). In a similar vein, the guidelines suggest sentences for car jacking (medians of about 11 years) that are double the preferences of respondents (median sentences of five years). For other crimes, particularly those where individuals may be injured or killed, respondents suggested sentences that were longer than those found in the guidelines (one interesting example of this was poisoning by tampering with over-the-counter drugs. The survey respondents suggested median sentences that were five years longer than those suggested in the federal guidelines). It is interesting that the authors find greater public concern over smuggling illegal immigrants and civil rights

violations than the federal guidelines would indicate. But here, as elsewhere in the study, the details of the crime matter for how respondents might sentence violators. Smuggling immigrants into the country for profit prompted sentences that were two to three years longer than those considered appropriate when family members were being smuggled in. Further, respondents were more concerned with vandalism to Catholic churches than with vandalism to Protestant churches or to Jewish or Muslim temples.

The second problem with the normative consensus argument is that there is significant disagreement among population groups about sentencing lengths for individuals convicted of federal crimes. For example, Rossi and Berk's analysis shows that African-Americans generally suggest less severe sentences than does the white population. African-Americans are also much less punitive when sentencing drug trafficking crimes and street crimes than are survey respondents generally. The length of suggested sentences also declines as education increases. In addition, respondents who live in the New England states appear to offer the least severe sentences, while those from the South tend to suggest the longest sentences overall. This hardly seems like evidence of consensus where criminal law is concerned.

The last, but perhaps most important, problem with Rossi and Berk's perspective is that it is not openly discussed until the last paragraph of the last chapter in the book. This is important because, as noted in the first chapter of the book, previous research has not dealt with how public opinion about criminal sentencing may be influenced by the media, political leaders, interest groups, or the actual practices of the courts. Rossi and Berk's research offers the potential to begin addressing these important issues. Unfortunately, the work remains narrowly focused on reporting the survey findings and rarely considers the mechanisms that shape public opinion. Thus, several times Rossi and Berk are forced to admit that they have no explanation for findings that do not fit their model of consensus. More important, they miss the opportunity to investigate the political implications of why groups such as African-Americans tend to be more lenient in their criminal sentencing opinions than the majority population.

In *Just Punishments* Rossi and Berk report their findings in a clear and readable style. It is their own findings that cast doubt on the interpretation they suggest. Many readers will find that the data can be interpreted within a Gramscian (or some other) framework as easily (and with greater satisfaction) as within the perspective offered by the authors.

Overlooking Nazareth: The Ethnography of Exclusion in Galilee. By Dan Rabinowitz. New York: Cambridge University Press, 1997. Pp. xiv+222. \$59.95 (cloth); \$19.95 (paper).

Rebecca L. Torstick
Indiana University, South Bend

Some commentators on Israel have noted that the true test for Israeli democracy has always been, not negotiations with a separate Palestinian political entity, but its treatment of its own Palestinian citizens. Dan Rabinowitz's ethnography of Jewish-Palestinian encounters in Natzeret Illit, a northern town intended to contain Arab Nazareth, comes at an opportune moment to test whether Israeli society would in fact pass muster. Rabinowitz has set out to explore "everyday exclusionism" (p. 11) in Israel by focusing on those areas where Jewish Israelis must share their landscape with a sizable population of Palestinian Israelis. In guiding our walk along this border, he explores a larger concern—how well does liberalism fare in political situations characterized by unequal relations among the state's citizenry? Do modern liberal states actually deal in a rational and even-handed manner with all their citizens, regardless of ethnic or minority status? Through a series of set pieces, from real-estate transactions and nursery schools to biographical sketches of three Palestinian men and a political campaign, Rabinowitz seeks to document the complex realities of life in this border zone and thus to expose how liberalism can produce an increasingly illiberal society.

In his first section, "Bigoted Liberals," Rabinowitz explores the limits of liberalism in Natzerat Illit by examining arenas where Jewish and Palestinian residents encounter one another. He focuses on how individual Jewish and Palestinian residents negotiate their encounters with members of the other community. How do Jewish residents explain selling/renting apartments to Palestinians, given the strong emphasis within Zionism on pioneering and settlement? How do Palestinian residents maneuver to their advantage through the Jewish-controlled educational sphere? In answering these questions, he did not always provide enough information. We never learn which Jewish Israelis were renting/selling apartments to Palestinians—was it the members of a particular social class, profession, ethnic or religious background? Rabinowitz glosses over divisions within Israeli Jewish society with the unfortunate consequence that his work cannot contribute to rupturing prevalent stereotypes about Mizrahi intolerance, although he himself wants to suggest that the perceived intolerance of the town's Jewish residents is a logical, if regrettable, outcome of their border zone location.

The second part of the book continues by examining the concept of resistance as it may usefully be applied to Palestinians living in Natzerat Illit. Specifically, he tries to demonstrate that the individual rational agency of particular Palestinians (a doctor, basketball coach, and candidate for political office) allows Jewish Israelis to come to "trust" them.

Such trust, however, is not generalized to the wider Palestinian collective, and so their agency does little to promote fairer treatment of Palestinians as a whole. This is not a new insight—a number of scholars have noted that positive encounters between individuals do not translate into improved intergroup relations.

One troubling area is Rabinowitz's contention that there are few ethnographic monographs on Palestinian Israelis (p. 17). There are a number of good recent works that are missing from his bibliography—Davida Wood's study of Arrabe ("The Boundless Courtyard: Palestinian Israelis and the Politics of Uncertainty" [Ph.D. dissertation, Princeton University, 1994]), Michael Romann and Alex Weingrod's work on Arab-Jewish interactions in Jerusalem (*Living together Separately: Arabs and Jews in Contemporary Jerusalem* [Princeton University Press, 1991]), Majid Al-Haj's study of Shefar-A'm (*Social Change and Family Processes: Arab Communities in Shefar A'm* [Westview Press, 1987]), and my own work on coexistence in Acre ("Raising and Rupturing Boundaries: The Politics of Identity in Acre, Israel" [Ph.D. dissertation, Washington University, 1993] and "Educating for Democracy in Israel: Combating or Perpetuating Racism?" *Identities: Global Studies in Culture and Power* [1 (1995): 367–90]). Is this further evidence of what Virginia Dominguez called the paradox of creating "oneness" in Israeli Jewish society—a process that produces a population that is "self-absorbed, uninterested in others, except when it perceives it must be interested in them—for reasons of security, international relations, trade or finance" (*People as Subject, People as Object: Selfhood and Peoplehood in Contemporary Israel* [University of Wisconsin Press, 1989, p. 153])?

While claiming to be sympathetic to the move in anthropology for reflexivity in ethnographic accounts, Rabinowitz still situates his own work within a corpus defined by male Israeli academics. At one level, he is careful to place himself in his field situation. We learn he is Ashkenazi, middle-class, a journalist; he performs his military reserve duty (although not in the Occupied Territories), has played professional basketball, and is sympathetic to the Palestinian plight. At another, he does not honestly confront and come to terms with his own privileged and powerful position vis-à-vis the residents of Natzerat Illit. Why should they have trusted him? If he wanted to explore the limits of liberalism in Israeli society, he could have studied the so-called liberal Israelis of North Tel Aviv.

In the end, we learn more about Dan Rabinowitz and the society that formed him than we do about the world of Jewish and Palestinian residents of Natzerat Illit. For this reason alone, this is an account worth reading if one wants to understand the very real difficulties in finding solutions to the Israeli-Palestinian issue. As Rabinowitz himself admits, he still takes the Israeli (Jewish) side of the question more seriously and less judgmentally (p. 189). Rather than having another version of life according to the dominant class, we need to hear from the Palestinian and Jewish Israelis whose voices have been silenced and marginalized. Rabinowitz is not the person to accomplish this.

Der Schweizerische Bauernkrieg von 1653: Politische Sozialgeschichte—Sozialgeschichte eines politischen Ereignisses. By Andreas Suter. Tübingen: Biblioteca Academica Verlag, 1997. Pp. 687.

Mark R. Forster
Connecticut College

Andreas Suter's study of the Swiss Peasant's War of 1653 is surely the definitive study of this seminal event in early modern Swiss history. This book goes, however, a step further. Moving beyond the narrow concerns of much German-language historical literature of this period, Suter successfully engages a whole range of important theoretical issues of concern to historians, sociologists, and anthropologists.

The subject of this book is not well known outside Switzerland. In 1653 peasants in the Entlebuch region, a district governed by the city of Luzern, protested a number of new taxes and monetary reforms undertaken by the authorities. Protest led quickly, in the terminology of 17th-century contemporaries, to open "revolt," then to a "revolutionary situation," and finally to a full-scale "peasant's war." The peasants developed a strong sense of unity and collective purpose, and the rebellion quickly spread to neighboring regions of Switzerland. Suter emphasizes that the peasants and their leaders (mostly landed peasants with considerable political experience) were determined, flexible, and capable of learning quickly in new situations.

Ultimately, however, like all early modern peasant revolts, this rebellion was suppressed by military force followed by a savage repression. As peasant resistance went underground, rebel leaders increasingly drew on Swiss historical memory, in particular the stories of the defeats of the Habsburgs in the 15th century and the semilegendary tale of William Tell, to construct a political discourse that focused on the tyranny of the authorities. This charged language laid the groundwork for the assassination, in September 1653, of several important Luzern officials on a lonely road in the Entlebuch, understood as a *Tyrannenmord* (tyrannicide) by the population. Suter argues that this assassination indicates a level of "treachery" (*Heimtücke*) and brutality that was unknown in other social conflicts in Switzerland in this period (p. 307).

Suter's primary goal in this book is to examine the relationship between long-term structures and historical events. How does a historian acknowledge and account for historical structures (economic trends, demographic changes, social structures, long-term cultural patterns) without predetermining the outcome of a particular event? In order to analyze the relationship between structure and event, Suter divides his book into two parts. Part 1 examines the Peasant's War from "the very short perspective of slow motion" (p. 45). Here he describes and analyzes events in close detail. Part 2 of the book examines long-term structural issues. Suter's organization consciously reverses that of Fernand Braudel's *The*

Mediterranean in the Age of Philip II (Harper and Row, 1972), giving primacy to events not structures.

Suter's discussion of events emphasizes the role of political leadership, focusing on the ways in which peasant leaders learned from and exploited events happening around them. Suter's extensive archival research shows that the rebels in 1653 made informed decisions, particularly at the decisive moments as they moved from protest to revolt to revolution to full-scale war against the authorities. Thus, although events had their own dynamic, people made real decisions. Suter also has a sophisticated understanding of the role of communication and language in peasant revolts. He details the role of economic ties and market connections in bringing Swiss peasants together as well as the use of traditions and historical memory in strengthening support for the rebellion.

Suter turns to a discussion of structural reasons behind the Peasant's War in part 2 of the book. Here the discussion is somewhat less original. The chapters discuss (1) the economic causes of the revolt, (2) the fiscal and political preconditions, (3) the role of "collective learning processes" and culture, and (4) the impact of lordship, economic ties, and social division on the "structuring" of the revolt. As this list of themes indicates, there are many important links between structure and the issues discussed in part 1. Perhaps most interesting is the significance Suter gives to cultural factors. He emphasizes here again the various ways peasants used cultural traditions, but also indicates the limitations imposed by tradition and historical memory. Thus the memory of major military victories in the 14th century gave peasants the willingness to revolt, but it also made them overconfident in their conflict with the authorities. Furthermore, because historical experience was passed down orally, it was "fragmented" and "inexact" and easily misled the rebels in 1653 (p. 455).

This beautifully produced volume is clearly organized with an English summary, illustrations, appendixes, and invaluable chapter summaries. Suter's scholarly apparatus demonstrates an impressive knowledge of the relevant literature in French and English as well as in German, an important factor in the quality of the book. At times, however, Suter has gone too far in his effort to be all inclusive. Although certain themes stand out (the "learning process" of the rebels and the culture structures within which they operated) the reader can get lost in the details. What does Suter believe were the most important factors, the most significant aspects of the process of rebellion in 1653? This weakness is more than compensated for by the discussions of theoretical and methodological issues, which are intelligent and thoughtful and based on an analysis of real events.

At the Crossroads: Mexican Migration and U.S. Policy. Edited by Frank D. Bean, Rodolfo O. De la Garza, Bryan R. Roberts, and Sidney Weintraub. Lanham, Md.: Rowman and Littlefield, 1997. Pp. viii+322. \$62.50 (cloth); \$23.95 (paper).

Jorge Durand
University of Guadalajara

Reviewing a book published by four analysts who write the introduction and conclusions together is an endeavor that may be redundant. What needs to be said about the work, they say clearly, well, and in few words. The first recommendation to a reader interested in knowing the impact of diverse public policies in Mexico and the United States and their relation to immigration would be to read the introduction and conclusions.

Once the table of contents offered in the introduction is known, the reader can choose different approaches. One approach can be made on the basis of countries. Five articles are focused on the United States, three on Mexico, and two on the situation on both sides of the border. Another approach is thematic, focused on four areas. First, there is the history of Mexican emigration and its relationship to changes in U.S. migratory policies, an issue that is repeated, perhaps more than necessary, in various chapters. A second topic refers to economic, social, and agricultural policies and their implications for migratory behavior. The third question has to do with the impact of Mexican migration in the United States in three areas: demographics, the tax system, and the job market. A fourth topic considers the course of each country's foreign policy in regard to immigration.

The book's aim is provocative: to interpret Mexican-U.S. migration in a way that differs from the traditional, well-known mode, based on the study of migrants' typical decisions and on the characteristics of their local surroundings, which are ultimately regional. The idea is that public policies—above all those dealing with employment and social services in the United States and development policies in Mexico—affect the volume of migration, which is a fundamental point of the discussion.

This aim is also the volume's weakness. The analysis of the impact of public policies remains on the national level quite often, based on global quantitative reports. This becomes a problem inasmuch as immigration is a markedly regional phenomenon, both in regard to the areas of migrants' origin and in regard to the areas of destination in the United States. However acute and sophisticated the analysis of Mexican information may be, it fails to take into account in a precise and efficient manner the migratory situation concentrated in certain regions, whose structures and social dynamics specifically embrace nationwide policies. Further, the very existence of migration is a factor of differentiation showing the manner in which diverse regions of Mexico process public policies. This evidence would have required a truly regional analysis of the impact of each public policy.

This is especially notorious in studies on Mexico. In defense of the authors, it must be said that a sharp contrast is noted between information sources at the state level in Mexico and the United States, which will surely one day be the cause of bilateral discussion. The existence of truly different jurisdictions in the United States assures the local gathering and processing of information, something that is still done irregularly and randomly in Mexico.

Thus, the volume detects not only social phenomena but also different platforms from which investigations are launched on both sides of the border. Therefore, one of this work's virtues is precisely that of making evident the necessity and urgency of defining problems and articulating solutions that involve both countries at different levels, that is, conceiving and integrating, in an efficient bilateral dynamic, issues so dissimilar yet so related as foreign policy, migration, trade and economic development, social services, and border relations, without forgetting the needs that arise from an unavoidable agenda of bilateral research.

The Meaning of Data: Open and Closed Evidential Cultures in the Search for Gravitational Waves¹

H. M. Collins

Cardiff University

The meaning of scientific "data" depends on the "evidential culture" of laboratories. Using transcripts of interviews and conversations with scientists, open and closed evidential cultures are analyzed under three dimensions. For example, an Italian laboratory's evidential collectivism and an American laboratory's evidential individualism are contrasted. In this case—the detection of gravitational radiation—evidential cultures are found to be homologous with institutional settings. The data interpretation of the long-standing small science is being influenced by the growing global dominance of a new big science. The interesting technique of "involuntary blinding" has been used to enforce a uniform approach.

MIT campus (Cambridge, Massachusetts), the "Laughing Seafoods" restaurant, December 1996.—Over dinner, two groups of experimenters are arguing about how gravitational radiation data should be interpreted. The data they are talking about are coincident bursts of energy affecting two or more widely separated resonant metal bars. The team from Frascati

¹ I am grateful to Ingemar Bohlin, Christa Knellwolf, Peter Dear, Dan Kennefick, Ian Welsh, and Robert Evans for helpful comments on various drafts of this paper. Many of the lively remarks of audiences at the Wellcome Institute in London, Cardiff University, Sussex University, and the University of New South Wales have also been incorporated. The constructive comments of *AJS* referees have influenced the text significantly. I am in debt to the many scientists who, over the years, have treated me as a colleague and have allowed me to share in a great scientific adventure spanning two millennia; I have learned much from their vigorous intellectual curiosity and refreshing optimism. All the respondents quoted at length read the paper and offered useful feedback, most of which has been incorporated; I do not thank them personally here so as to maintain a semblance of anonymity. Two physicists not quoted at length, Sam Finn and Peter Saulson, read a draft of the paper with enormous care and gave me encouragement and advice. I am, of course, entirely responsible for the paper's faults. This research received generous financial support from the U.K. Economic and Social Research Council (grants R000235603 and R000236826). Correspondence may be directed to H. M. Collins, KES, Cardiff University, 50 Park Place, Cardiff, CF1 3AT, United Kingdom. E-mail: collinshm@cardiff.ac.uk

has suggested that it is reasonable to publish reports that say that coincidences have been seen, even if they do not claim that they are gravitational waves. The group from Louisiana State University (LSU) disagrees and would prefer to publish nothing.

LSU2: That's a difficult thing—it's not the—it's a reasonable thing to do, but it's not what most physicists want to see. Normally, you get to the end of an experiment and you give a result. "I didn't find the X-particle, and my results are consistent with 'X-particles will not be produced this way because their cross section must be less than something.'" That's a negative result, and a positive result is "I saw the thing. . . ."

If you say, "I saw something very unusual, that doesn't seem as though it could be due to chance," then, what do you do—that's the grey area—you're not making a negative claim, you're not making a positive claim, you're somehow in the middle.

Frascati1: You know, one consequence of all this is the following—[it is] very strange—that you do a coincidence experiment; if you find nothing—in other words, if you find a number of coincidences equal to the number of accidentals—then you publish, and you give the upper limit [that is, you say what the maximum flux of gravity waves can be]. If you find an excess [a positive result], you don't publish—that's the conclusion [laughter]. Don't you think so?

LSU1: There is a danger of that—there is.

Frascati1: If you find nothing, you publish; if you find an excess, you don't publish—I mean an excess with minimum probability of 1% or 1 per 100. You see, the only excuse not to publish is if one is afraid that by publishing there might be bad consequences, like not getting money from the funding agency, or things of that sort. Then you can be excused for not publishing, but it is a bad thing. But—of course . . .

LSU1: Well, we have been in that situation too—in years past—I don't think we are anymore. But—there's . . .

LSU2: The problem is particularly difficult for us. It's not the normal problem in science, I don't think—frankly. Or at least, a fair fraction of science. If we really did see a gravitational wave at our current level of sensitivity, it would be in conflict with what many people believe is well-founded, established facts. It might be in conflict with the general theory of relativity, which is not . . .

Frascati1: Nol . . .

LSU2: Well, if we saw gravity waves, what possible conclusions can you draw? What is the theory? . . .

Frascati1: Well, first point is that a coincidence excess might be due to another phenomenon, not gravitational waves—a very important phenomenon, even, at this point, even more important, that nobody knows about.

LSU2: That's a second possibility, that there's unanticipated physics operating. Well, that would be—that's a good reason to publish if you've got evidence that's pretty clear. But in my opinion, you have to be pretty certain that the chance of this being ordinary accidental behavior has to be pretty small to make it worth it to even start to believe such a thing.

Frascati1: But the definition of small is not easy. . . .

Frascati1: I think each one follows his own philosophy. My personal phi-

losophy is the following: if I have a situation where I don't know whether I should do something or not to do it, I do it. Because the outcome is better than no outcome. Now this does not apply . . .

LSU1: I was gonna' say that that could lead to overpopulation of the world. [laughter]

Frascati1: Because, if you don't do it, it's nothing—if you do it, it's something.¹

INTRODUCTION

The above conversation represents a clash between open and closed "evidential cultures." The evidential cultures are related to the institutional contexts in which the two laboratories find themselves. To understand the passage sociologically, one needs to know something of the history of the field of physics in which the laboratories work, something of recent events in that field, and something of physics itself. In this kind of study, one has to know a little science in order to understand the sociology, and this means the structure of the article presents a chicken-and-egg problem: Should it be physics first or sociology first? I have tried to solve the problem in the following way:

Following this introduction, I briefly compare sociology of scientific knowledge (SSK), the approach used in this article, and earlier sociology of science. I indicate the methodological leaning of the article and the scope of the case study.

The next section contains a first description of gravitational radiation research and an indication of the features of the field that makes it an especially interesting site for sociological research. After that, I introduce the idea of evidential cultures and illustrate it with examples and quotations covering publication strategy. I then set the evidential cultures in their institutional context, introducing a little more physics on the way.

The second main section of the article introduces two recent incidents in the field. It begins, perforce, with more physics, concentrating on data analysis. It moves on to discuss the solution to the clash of cultures that has been put in place by the American laboratory against the wishes of the Italian laboratory.

A description of a recent controversial claim to have detected gravita-

¹ With the exception of a few E-mails, all the quotations in the paper are taken from interviews or conversations recorded by the author. All have been cleared for publication by the speakers. Most respondents felt more comfortable with the anonymous "style" that has been adopted throughout. Of course, most insiders will easily recognize many speakers, but occasionally, because I was asked, or because I thought it more appropriate, I have disguised identity more thoroughly. I retain a master copy of the paper with a complete key. Several respondents asked me to edit their "uhms," "ers," "you knows," and "sort ofs," and I have removed them throughout.

tional waves brings us back to the Laughing Seafoods conversation. The article ends with the conclusion that the meaning of data in physics depends on evidential culture and institutional setting.

APPROACH AND METHOD

The sociology of scientific knowledge (SSK), the style of analysis adopted in this article, is a quarter of a century old. Nevertheless, it is still frequently referred to as a new and controversial approach. Compared to the earlier, Mertonian tradition, it differs in several ways. It does not start with the assumption that there is a single epistemologically efficacious set of social norms that guides scientists to the truth. Though this article could be said to follow in the Mertonian tradition in the sense of being concerned with patterns of scientific action, it compares the actions of one group of scientists with the rather different actions of another group, without making a judgment about which is better.

It might be said that even in this the approach follows Merton, for in his paper on ambivalence (Merton 1963), he discusses the tension between the desire for priority in science and the imperative to be right. But SSK, though it describes individual dilemmas and strains within patterns of action, is primarily concerned with groups, their social settings, and their scientific cultures, or "forms-of-life" (Wittgenstein 1953; Winch 1958). Moreover, systematic differences between groups are shown to affect what counts as scientific knowledge. The analyst has to enter far more deeply into the science than was necessary in the Mertonian tradition, sometimes embracing competing scientific positions. To do this, views about scientific truth have to be set aside, at least in respect of the matter under examination and at least temporarily.³

³ For sociological ambivalence, see Merton (1963). The tradition of the analysis of the norms of science goes back at least as far as Merton (1942) (See also Merton 1973; Shapin 1994). The tradition of empirical and historical studies of the SSK type goes back at least as far as Ludwik Fleck, though it was not recognized as a tradition until much later and some of the contributors listed below might resist the label. Early works include Fleck ([1935] 1979), Kuhn (1961, 1962), Collins (1975, 1981a, 1992), Holton (1978), Latour and Woolgar (1979), Shapin (1979), Knorr-Cetina (1981), MacKenzie (1981), Pickering (1981, 1984), Pinch (1981, 1986), Travis (1981), Collins and Pinch (1982), Glaser (1983), Lynch (1985), Galison (1987), Shapin and Schaffer (1987). Within this later tradition, this article stresses the way networks of scientists and of laboratories negotiate the meaning of data between themselves. This style of research goes back at least as far as the study reported in Collins (1975). While adopting "methodological relativism" (Collins 1981c) and the principle of symmetry (Bloor 1973), it resists the extension of symmetry to the relationship between human and nonhuman entities (Callon and Latour 1992; Collins and Yearley 1992). For easily accessible reviews of the field written from the perspective of more recent trends, see articles in the *Annual Review of Sociology* by Collins (1983a) and Shapin (1995).

This setting aside of the scientific truth has led some commentators to imagine that SSK endorses irrationality or some such, but, on the contrary, it differs from other ways of analyzing science in imputing *more* rationality to scientists than they allow themselves. When a minority of scientists hold scientific views that most other scientists reject, the former are often taken to be acting irrationally or unreasonably. SSK, however, with its "symmetrical" approach (Bloor 1973, 1976), tries to recapture the rationale of all sides; the mistake is to think that rationality can lead in only one direction (Franklin 1994; Collins 1994). One might say that SSK tries to make sense where scientists—and those sociologists who accept the majority's interpretation of matters scientific—make nonsense (of the minority view). That is a difference between sociology of scientific knowledge and science (and the older sociological tradition) that many have misunderstood. It arises because it is scientists' responsibility to make science and, therefore, to make the nonsense with which the sense can be contrasted; it is not sociologists' responsibility to make either.

To carry through this kind of analysis, the sociologist must adopt a stance widely known as "methodological relativism." Methodological relativism implies that, for the purpose of analysis, the sociologist "brackets out" received views of scientific sense and nonsense. Instead, the focus is on the interpretative flexibility that is available to scientists in respect to their findings; the networks of beliefs, practices, and interests that favor one interpretation over another; and, ideally, the way that one interpretation rather than another comes to predominate.⁴ If received views about scientific truth are not set aside, the whole enterprise becomes biased at best or circular at worst. Whether or not this methodological relativism implies a commitment to an epistemological or ontological position has been the topic of much philosophical debate, which need not concern us here. Here, all that we need in the way of epistemology and ontology is some slack.

The more narrow methodological stance adopted in this article is "participant comprehension" (Collins 1984). Participant comprehension is an interpretation of participant observation under which the field-worker tries to acquire as high a degree of native competence as possible and interaction is maximized without worrying about disturbing the field site; this ideal should always direct the research effort, even though the degree of native competence attained will vary from study to study. In this case, my understanding of the field of gravitational radiation detection has been gained from discussions between 1971 and 1976 with most of the major actors in the field and from more intense contact with the field since 1993.

⁴ The approach has been referred to as the empirical program of relativism, or EPOR (Collins 1981b, 1983b)

Recent work has been made possible by financial support covering the whole of my time in the year 1995–96 and most of my time from 1996 to 2001. During these periods, I have, so far, carried out around 150 tape-recorded interviews in the United States, Britain, Germany, Italy, and Australia; attended around a dozen conferences and committee meetings (including private meetings at funding agencies); engaged in scores of E-mail and telephone interchanges with physicists; spent periods in laboratories and in on-site visits throughout the world; and, perhaps most important, joked, chatted, and argued with physicists about their work and mine, in corridors, cafés, cars, buses, restaurants, and a boat.⁵ Had my early papers not been seen as reasonably accurate and responsible portrayals of events, I would not have been able to reenter the field in the 1990s. That said, I would not claim to have achieved anything like full native competence in gravitational radiation research—I cannot actually do it myself—but I believe I have gained enough understanding to be able to carry out the kind of sociological analysis presented here.⁶ On two memorable occasions I have even managed to argue my corner while talking minor matters of gravitational wave physics with physicists, and on a third occasion I was told that my critical queries would affect the design of a major new instrument.

This article has been seen by the principle characters discussed within it and by other sympathetic physicist informants, and they have found no fault with the physics. Needless to say, some respondents would have preferred the analysis to be less neutral, and this is always likely given that explaining one side's position looks like attacking the other's. Other physicists commended the paper as an accurate appreciation of a recent episode in the history of their enterprise. The opinions offered by scientists

⁵ The places I have visited once or many times so far in connection with this project include (number of visits in brackets) the universities of Bristol (3+), Reading, Glasgow (2+), Sussex, Cardiff (2+), Hannover (2), Leiden, Maryland (5+), MIT (3+), Irvine, Stanford (3), Caltech (5+), Rochester (3), Louisiana State (5+), Tor Vergata-Rome, Padua, Milan, Perth, Adelaide, Canberra. I have also been to industrial labs, including IBM (3), Bell Labs (3), Hughes Aircraft, CSIRO Sydney, NASA-Goddard Space Center. Other locations include the Max Planck Institute—Munich, Sophia-Antipolis, the LIGO site at Hanford Nuclear Reservation (2), the LIGO site at Livingston, Louisiana (2), the GEO600 site near Hannover (2), the VIRGO site near Pisa (2), the Frascati labs (3+), the NSF in Washington (4+); the offices of April Burke and Kevin Kelly in Washington; the homes of Joel Sinsky in Baltimore, David Zipoy in Punta Gorda, Florida, and Bob Forward near Inverness. I have also attended conferences or committee meetings in Washington, Pisa, Orsay-Paris, Jerusalem, Geneva, Boston, Hanford, Livingston, and the California Institute of Technology (Caltech).

⁶ In this judgment, I benefit from the experience of other case studies, including one, on parapsychology, in which I became a full-blown expert, and another, on the theory of amorphous semiconductors, which I had to abandon because I could understand none of the science.

on the merits of the sociological analysis were not correlated with their position in the scientific debate.

Participant comprehension licenses a certain approach to interview material. Interview extracts *illustrate* what analysts understand to be going on as a result of their experiences in the field and are used to convey this understanding to a less expert audience; they are not thought of as *data*.⁷ The scientific warrant of this paper as a whole is its internal logic and the fieldwork it reports.

There is an interesting tension between methodological relativism and participant comprehension, because few native physicists are relativists. An acute problem arises when the analyst looks at a scientific controversy, because then the job is to acquire native comprehension of both sides of a debate—something that natives rarely need. One might express this as an issue of methodology. One of my failings as a participant-observer/comprehender is that I cannot generate the levels of disdain for physicists that one physicist can generate for another; I know what I am missing because, given that I have no commitment to symmetry in my own field, I can readily generate such levels of disdain for some of my fellow social scientists.⁸ Fortunately, given the qualifications and compromises described above, science does not have to be perfect to be science; it is fortunate, because otherwise there would be no social science nor any physical science either.

THE DETECTION OF GRAVITATIONAL RADIATION

According to the General Theory of Relativity, gravitational radiation is emitted when masses move. Because they are so weak, only huge cosmic catastrophes, such as the explosion or collision of stars, can generate detectable fluxes of gravitational waves. If the shape of gravitational waves is ever revealed in detail, a new field of astronomy will come into existence; gravitational radiation is the only direct way to see into black holes and many other features of the universe. The better part of a billion dollars is now being spent worldwide in order to realize the aim of a new gravitational astronomy using the technique of "interferometry." At the same time, the older and cheaper program of detection, pioneered in the 1960s by Joseph Weber of the University of Maryland and using resonant metal

⁷ Compare the notion of interview material found in Mulkay, Potter, and Yearley (1983) and the rebuttal in Collins (1983b).

⁸ The classic participant-observer study (Festinger et al. 1956) was flawed for the opposite reason: the "participants" had too much disdain for their subjects. Festinger et al. were not trying to "comprehend" but only to observe the "irrational" objects of their study (Collins 1984).

bars, continues with instruments of increasing sensitivity. In this technique, a gravitational wave would make its presence felt as a coincident input of energy on two or more bars separated by thousands of miles.

In the early 1970s, Weber announced that he had seen gravitational waves with his bars. In earlier papers, I reported on the controversy that followed (Collins 1975, 1981a, [1985] 1992). Weber's results were widely disbelieved, and by 1975 they had few supporters. The episode is an important part of the background to the story that follows.

Gravitational radiation detectors, like many pathbreaking instruments, have a very poor "signal to noise ratio." According to the dominant theories, in spite of improvements in sensitivity by a factor of perhaps a million over the last four decades, none of the detectors that will be completed before around the year 2006 ought to be able to see the effect unless they are very lucky.⁹

The estimate of an instrument's likelihood of seeing gravitational waves is based on calculations of its sensitivity, the gravitational energy that ought to be emitted during cosmic catastrophes—such as the spiralling-in and coalescence of neutron stars in orbit round each other—and the likely frequency of such events at various distances from us in the heavens. The opportunity for a lucky break arises because one or two such events could take place relatively close to us, in our own or a nearby galaxy, giving a unusually strong signal at the surface of the earth.¹⁰ Even with luck, however, it is difficult to tell when a disturbance in a detector is a signal and when it is just a peak of accidental noise, because a "strong" gravitational wave is still about the weakest thing to be measured since measurement began.

It is because gravitational waves are so weak that to separate them from noise in the detector it is necessary as a first step to look for coincident signals on two or more widely separated detectors. There is one research group that runs two gravitational antennae that are widely separated (in Rome and Geneva), but the credibility of a finding is boosted if the coincidences are seen by apparatuses belonging to separate laboratories. In effect, since the middle 1970s, a device for detecting bursts of gravitational waves has consisted of two or more antennae run by two or more groups, most often located in two or more countries.¹¹

⁹ *Pace* the theory of bars developed by Weber and Giuliano Preparata of the University of Milan (Weber 1984; Preparata 1990).

¹⁰ On the other hand, we might be wrong about the constitution of the heavens (see below) and each new generation of detectors might just turn out to have crossed the crucial sensitivity barrier for as yet unknown events and processes.

¹¹ This is true only for signals that come in bursts. Detection of the waves from continuous sources, such as those from pulsars, can, in principle, be accomplished with only one detector. The random background gravitational radiation left over from the big

Collaboration and Conflict

When science is done by individuals or by small teams, experiments are worked up, data are taken, analyzed, and interpreted, and then "results" are published, all under the control of one laboratory. The boundary between the private world of the laboratory and the public world of journals and conferences can be policed by scientists working on their own or leading a physically and morally integrated team to a consensus. Of course, the wider world of science might not accept the results, and controversy and rejection might follow, as in the case of Weber's early findings. Nevertheless, the differences between developing an experiment, evaluating the data, and announcing the results for further evaluation beyond the laboratory are clear. This is important, because it is our distance from the untidy world of the laboratory, with its continual adjustments, choices, and need for active interpretation, that helps separate science from mundane activities. The idea of the laboratory as a place where the consequences of nature's agency are observed and reported with a minimum of human intervention is maintained by keeping the laboratory closed.¹² But, in the case of gravitational radiation data, because findings are sets of coincident readings, responsibility for initial interpretation is the prerogative of more than one team. All science is, in the last resort, communal, but in the search for bursts of gravitational waves, international collaboration is a *condition* for the *initial* announcement of a credible result.

Such an intrinsically intimate relationship between diverse groups of scientists is very unusual. There are, of course, many examples of replication by other groups or collaborations in which the findings of different

bang or produced by the huge numbers of collapses and explosions taking place throughout the universe needs two detectors that are close together. Such results would still need to be replicated if they were to be believed, but they are not intrinsically "coincidence experiments." Until very recently, the history of the search for gravitational radiation was the history of the search for bursts of energy, and here, when I refer to gravitational radiation, I will be referring to bursts. All the incidents reported here concern the search for bursts

¹² This is a very important principle in the history and sociology of scientific knowledge. It enables one to establish the difference between an "experiment," a "result" announced in public, and a "demonstration," which is a display of well-established results, perhaps using techniques to enhance the effect that would be considered to be cheating in an experiment. (See, e.g., Gooding 1985; Shapin and Schaffer 1987; Collins 1988). Notoriously, historians and sociologists of scientific knowledge invade the privacy of the laboratory and threaten these distinctions. Holton's famous study of Millikan's oil drop experiment and many other historical and sociological analyses of famous experiments (see n. 3 above) exhibit the problems. This very article may cause discomfort because of the way it exposes what would normally be the private thoughts and conversations of scientists, not yet honed and refined for public consumption. I am sorry that my job sometimes makes me appear to trespass upon the conventions of friendship.

laboratories are aggregated, but in this case, there are no findings at all until the raw outputs of the groups are combined. Because of this, clashing research styles will cause the laboratories to pull against each other like convicts from a chain gang trying to run in opposite directions. As with the convicts, once one knows what to look for, the predicament is striking, giving the analyst a great advantage.

There is no reason to think that this forced grinding together of potentially different scientific cultures makes for untypical science in any other way. The three dimensions of evidential culture to be discussed below are likely to be found across the sciences, wherever conditions are appropriate, even though they may be less obvious in places where different laboratories are isolated within separate research programs. It remains to be seen, however, if divergencies in evidential culture are typical only of relatively small sciences. Big sciences, such as high-energy physics and gravitational wave interferometry (see below), may or may not be able to sustain a variety of styles in spite of being global enterprises.

THREE DIMENSIONS OF EVIDENTIAL CULTURE

The LSU group and Frascati group are chained together by their joint research on gravitational radiation. As revealed in the Laughing Seafoods conversation, their evidential cultures do pull in opposite directions. I am going to suggest that evidential culture has three dimensions: "evidential collectivism" versus "evidential individualism"; high versus low evidential significance; and high versus low evidential thresholds.

Evidential Collectivism and Evidential Individualism

Evidential individualists believe that it is the job of the individual or individual laboratory to take responsibility for the validity and meaning of scientific results. Under this philosophy, as much responsibility as possible is gathered into the individual or individual laboratory. In contrast, evidential collectivists believe that it is the job of the scientific collective to assess results from an early stage. In the case of both individualists and collectivists, the ultimate arbiter is, of course, the scientific community, but the individualist will consider it a matter of professional failure if the community rejects results sent forth from the laboratory while the collectivist thinks discussion by the core-set (Collins 1992)—whether the outcome is acceptance, rejection, or a demand for further clarification—to be a normal part of the scientific process.¹³

¹³ Evidential collectivism differs from the norm of "organized skepticism" (Merton 1942). Organized skepticism is a collective activity, but it is meant to *check* the validity of *individuals'* findings rather than interpret them. In any case, the idea of organized

The two philosophies can be usefully compared to styles of car driving.¹⁴ The first duty of a British driver, or an American driver is not to upset the equanimity of others who are driving acceptably—this is an individualist ethos. Rome, on the other hand, represents a case of driver collectivism. For example, on Rome's crowded ring-road, I once stopped my car on a busy roundabout while I conspicuously consulted a map. In Britain, such behavior would be almost unthinkable, and every passing driver would hoot and gesticulate, exhibiting what has become known as "road rage." In Rome, the approaching traffic simply skirted around the obstacle I presented without a second glance. Another driving example was offered to me by an American physicist. He explained that once, in Rome, he was parallel parked nose to tail on a road with a dense and slow-moving stream of traffic passing close to the parked cars. But, as his initial maneuvers indicated his desire to leave the parking space, the traffic squeezed itself across the road in such a way as to leave him room to extricate himself and join the traffic stream. The initial act of parking was, one might say, antisocial, illustrating, as he initially thought, the careless individualism of Roman drivers, but the cooperation of the community of drivers made it into a reasonable act.

In sum, in this part of Italy, responsibility for avoiding accidents and ensuring a smooth flow of traffic is passed to the community of drivers; this works well, without engendering fury. In Britain and America, the responsibility remains with the individual, and departures from the norms are met with sanctions.¹⁵

It has been put to me that the terms "individualism" and "collectivism" are here being used the wrong way round, and it is the Roman drivers who are the irresponsible individualists, while the American and British drivers are far less antisocial. But there is no such thing as a society made up entirely of individualists. The deep point is that in Rome it is the com-

skepticism has its roots in the supposed ready reproducibility of scientific findings rather than the intentionality of scientists' actions. For a discussion of the complexity of replication, understood through the notion of the "experimenters' regress," see Collins (1992). The Mertonian norm of "communism" refers to the collective *ownership* of results, not the collective establishment of results.

¹⁴ I am not suggesting that the boundaries of the collectivist driver culture coincide with the boundaries of evidential collectivism in science, merely that this is a similar set of social relationships found in a different field of activity.

¹⁵ That this is a matter of driving norms rather than national character can be seen because the relationships between car drivers and pedestrians is reversed in Italy and America. In most of America, drivers take it as their responsibility to avoid damaging pedestrians, who seem to think it their right to cause any amount of inconvenience to the driver. In Britain, it is the pedestrian who always defers to the car on pain of injury or, at least, indignation, and the same applies in Italy.

munity that is given responsibility for maintaining order, and that is why the individual can act with what looks like carelessness; the collectivity "repairs" any potential disruption before it becomes serious. In the British-American approach, it is individuals who are meant to retain responsibility for the smooth organization of traffic society.

These alternative ways of organizing society in respect of car driving reflect similar choices made throughout social and political life; after all, what we are talking about is different ways of dividing labor—in this case, it just happens to be cognitive labor. Consider that there are two ways to organize an undergraduate course in sociology during periods of high politicization of the subject. One can insist that every teacher of sociology presents an unbiased course, so that if, for example, they favor a Marxist approach, they also put the counterarguments, or one can allow each teacher to teach according to his or her biases but make sure that the faculty as a whole is balanced. The first of these solutions is individualist, the second is collectivist in the sense I am using the terms here. Consider again that those who insist that reflexivity is a vital part of the sociological analysis of scientific knowledge are individualists in the sense used here, because they believe that it is the duty of the *individual* to produce a complete analysis: the analysis must include not only a discussion of the social influences on the science under examination but also an analysis of the social influences on the analyst. The collectivist (such as myself, in this instance) believes it is satisfactory to complete the analysis of the social influences on the science, leaving other members of the community to analyze, if they are interested, the social influences on me (as in Ashmore 1989).

To return to gravitational radiation, I am arguing that the Frascatian approach (shared with the laboratory in Perth, Australia, to be introduced below) is set within the kind of collective ethos that is found on roads in and around Rome. The Louisianian approach to science is better compared to the driving style found on the majority of British and American roads.

Sociology would be much easier if only social life arranged itself as neatly as the analyst would like. Unfortunately, evidential collectivism is not the uniform style of Italian science or Australian science nor even any single laboratory within those countries. For example, the Frascati laboratory includes many members who are evidential individualists, and the discussion about whether to publish the contentious findings has been carried forward as much within the Frascati laboratory as between it and the rest of the world. For this reason, I will henceforth refer to the subset of Frascatians who do share the evidential culture under discussion as the "Frascati Team," while retaining the usage "Frascati group" for the

whole laboratory. What is more, there may well be groups of evidential collectivists scattered within the dominant culture of evidential individualism of Britain and America.

In spite of these reservations, it remains intriguing that the principle Italian physics journal, *Il Nuovo Cimento*, seems to express a collectivist philosophy. I asked its editor about his policy in respect to a recently published and notably "adventurous" paper. Joseph Weber (Weber and Radak 1996) claimed that his gravity wave data correlated with "gamma ray bursters." (The existence of this paper, I might add, was unknown to the very large majority of American gravitational radiation physicists, who generally no longer read *Il Nuovo Cimento* and who, in any case, gain most of their information from the informal networks within which Weber no longer has a voice.) The editor said:

Well, this is a complex issue, and it is very much a matter of the judgment of the editor in this case. . . . My opinion as an editor is that not all the papers have to be necessarily correct, but we must make every effort possible to publish papers which are not wrong—a priori wrong; if a paper is [obviously] wrong and we publish, this is unacceptable. But if a paper is not obviously wrong, and it is on a topic which is important, I think it can stimulate discussion. And the important thing is finally to have an answer—to promote the discussion and to reach a definite answer.

Now, in this sense, even if something is not absolutely proved to be final, it could generate other people to look, in order to disprove it, for example—to stimulate a search. . . .

This is a general approach that I have been following. There are, for example, other cases of theories which are not settled, which I decided to publish anyway in order to provoke—to generate discussion. . . .

So long as things are not done randomly, but are done by an editor in order to converge to an answer, that is OK.

While I will deal only with the contemporary scene, it is tempting to speculate about the historical roots of these phenomena. Steven Shapin's *Social History of Truth* (1994) examines the origins of scientists' actions in the norms of English gentlemanly behavior in the 17th century. The norms described by Shapin can, with a little stretching, be mapped onto evidential individualism. Questioning another gentleman's assertions could be construed as branding him a liar—a charge that could lead to a duel. The prohibition on "giving the lie" imposed a reciprocal obligation, making the validity of observations very much the responsibility of the individual, even while obeisance was given to the idea that the community was the ultimate arbiter. Thus, it may be that in exploring the relationship between English gentlemanly norms and science, Shapin has discovered the origins of evidential individualism and that other national traditions

or religious cultures may have given rise to different expectations of the way scientists should act, the traces of which are also still visible in the forms of life of modern science.¹⁶

Evidential Significance

Pinch (1981, 1986) was the first to discuss a component of what I am calling evidential culture. He described the search for solar neutrinos. As Pinch explained, experimenters try to detect neutrinos through their ability to transmute the chlorine atoms in a tank of cleaning fluid into radioactive argon. Putatively, radioactive argon atoms betray themselves as marks on a chart recorder. The experimenters can report this as "marks on a chart" (which is the only thing they see directly), "radioactive argon," "solar neutrinos," or other phenomena at different levels of generality. The higher the evidential significance, that is the longer the chain of inference, the more important the findings. But the more important the claim, the greater the risk of engendering opposition, and the longer the chain of inference, the more ways there are of being wrong. Thus, a choice of high evidential significance entails what we might call high "interpretative risk," while claiming low evidential significance involves only low interpretative risk.

In a similar manner, in gravitational radiation detection, the same coincidences can be reported simply as "coincidences" or as "gravitational waves." A claim to have seen gravitational waves is far harder to support and far less likely to be credible than a claim to have seen some coincidences. In the Laughing Seafoods conversation, we can see that the Frascati Team is happy to announce unadorned "coincidences," whereas the Louisiana group wants to announce nothing unless they are sure they have gravitational waves.

Evidential Threshold

A third independent dimension of evidential culture is choice of evidential threshold and concerns "statistical risk." Irrespective of whether the collectivity or the individual is considered to be the proper locus of scientific findings and irrespective of the degree of interpretative risk endorsed, a scientific culture might be more or less risk averse in terms of the level of certainty in the data that is taken to merit announcement or publication.

¹⁶ There is evidence that the same situation applied at one time in respect to England and Germany. Infeld in his autobiography (1941, p. 190) remarks that the attitude of prewar English journals was "better no paper than a wrong paper," while German journals felt "better a wrong paper than no paper at all."

The choice might be as simple as whether journals are full of two- or three-standard deviation results, as in the social sciences, or seven- or eight-standard deviation results or better, as in high-energy physics.¹⁷ In our language, the social sciences have a low evidential threshold, while high-energy physics has a high evidential threshold.

One might argue that the whole field of gravitational radiation research would never have started were it not for Joe Weber's high-risk—in both interpretative and statistical senses—reporting of putative phenomena. Weber may now be largely thought to have been wrong, but most scientists with a long acquaintance with gravitational wave physics would agree that he founded a field as a result of pressing forward obstinately with his claims.

In the Laughing Seafoods conversation and in the section on signal thresholds to be discussed below, one sees clearly that the Frascati Team endorses a strategy that involves high statistical risk—that is, low evidential thresholds—together with low interpretative risk (low evidential significance). The Louisiana group reveals the opposite position.

OPEN AND CLOSED EVIDENTIAL CULTURES

Putting together these three dimensions enables us to imagine an “evidential culture space.” In figure 1, the Louisianians are at one corner of the imaginary space—low collectivism, high significance, and high threshold, with the Frascati Team at the opposite corner—high collectivism, low significance, and low threshold.

What all three dimensions have in common is that the position adopted by the Frascatians on all of them tends to early release of relatively unprocessed data, whereas the position adopted by the Louisianians causes them to restrict access to their results until they have been much more highly processed. The Frascatians have an open evidential culture; the Louisianians have a closed evidential culture.

These two corners of evidential culture space also represent the change from the early days of gravitational radiation detection—when it was reasonable to “point a detector at the sky” and report what was seen—to the present—when what it is legitimate to see is very much more constrained by theoretical considerations (see below). The tension between the corners of the evidential culture space was experienced within the Frascati laboratory itself, and the positions there were described to me as the symptoms of, respectively, an *experimental animus* and a *mathematical animus*.

¹⁷ For a comparison of the “epistemic cultures” of different disciplines, see Knorr-Cetina (1991).

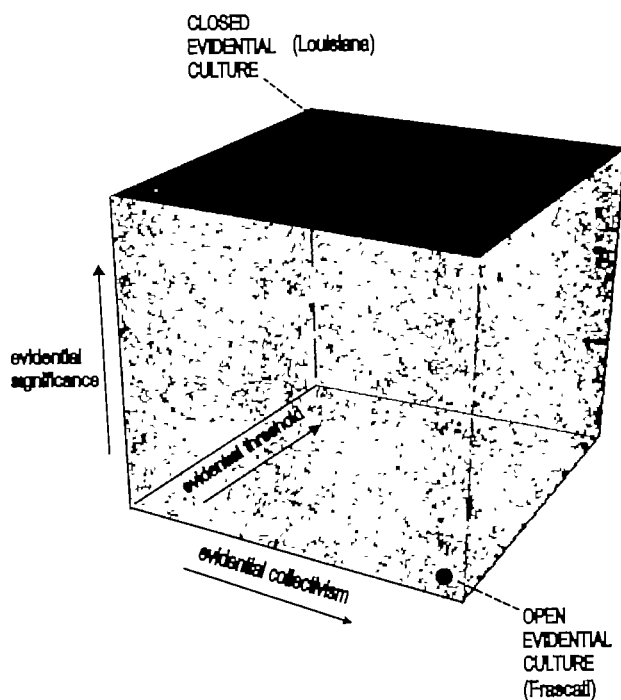


FIG. 1.—Evidential culture space

These are useful labels in this case because of the current role of theory in holding back speculation.¹⁸

The experimental animus is very well illustrated by the notorious 1989 publication (Aglietta et al.), which claimed that the room-temperature resonant bars run by Weber in Maryland and the Frascati Team near Rome found coincident signals, which were themselves in coincidence with neutrinos emitted by the 1987 visible supernova. In that paper, it was made clear that the energy level required to produce the signals was equivalent to the complete conversion into energy of a mass equal to 2,400 suns—something that even the members of the Frascati Team considered well beyond belief.¹⁹ At the same time, the levels of statistical significance re-

¹⁸ In a field such as biology, it might well be experiment that holds back "irresponsible" theorizing, so the terms "mathematical animus" and "experimental animus" would label different points in the same space.

¹⁹ But Joe Weber, Giuliano Preparata, and their colleagues would consider the claim reasonable, because under their theory of the bars' enhanced cross-section, only a fraction of a solar mass was required to produce the effect they apparently saw.

ported in the paper were not such as to avoid the charge of statistical massage. Nevertheless, the Frascati Team considered that this data should be published and looked at by the rest of the community; as far as they were concerned, something had been found, and it could not be wished away. A mathematical animus sees experiment as the servant of theory; it would restrict the presentation of data to that which makes sense in theoretical terms.

Again, the Frascati Team's evidential culture is well set out in the attitude to publication expressed in the following quotations recorded on separate occasions in 1996. Both reflect a leaning toward low evidential significance, the first combining it with high evidential collectivism, the second with readiness to adopt a low evidential threshold.

Frascati1: You should say "look"—no, it's a difference of how you publish. You should say "look, we found this result—we inform you—so if it's the case that you find the same results, you will report it to us." You should not say that you found gravitational waves. . . . Other people should say that.

Frascati2: There is a coincidence excess over what is expected. Now, is this coincidence excess enough to make an extraordinary claim or not? What is the probability that the excess is by chance? Calculating this, it is a few percent I understand. Is a few percent enough or not? Well maybe it is not enough to conjecture that these are gravitational waves, because to do that, you need extraordinary evidence. But it is enough to publish the results saying, honestly, "We have these results."

In contrast to the Frascati Team, a member of the LSU group provided the following by way of explanation of their unwillingness to be associated with anything but the most secure results:

LSU1: On thinking about it, I'm pretty sure that foremost in my mind was there is no way that I want to do anything that is going to have me labeled as a crackpot. . . . I suspect that underlying all of this is the desire to really say, "Well, OK, if my work is going to be significant, I've got to make sure that no one thinks that it's crazy."²⁰

The low evidential significance strategy would be fine if the development phase of an experiment could be kept clearly separate from the process of making a claim about the existence of new phenomena. As a member of another sympathetic group put it, "It is only by poking around at the data that you sometimes manage to make a discovery. Because of reproducibility, results must be the same or similar in the next experiment—and because of the ability to repeat the experiments . . . data massage is not as easy as [some] like to claim." But keeping report and interpretation

²⁰ We should read what is said here as an indicator of what counts as "crackpot behavior" or "craziness" in the respondent's reference group.

distinct are not easy; the meaning of a paper is in the eye of the beholder. It is not always the case that papers are read or interpreted in the same way as the author intends:

LSU: You were telling me that the paper you wrote with Weber, you did not claim gravitational waves. Everyone thought you did but . . .

Frascati1: Well, at CERN they said, "Oh—you publish a paper with Weber—Oh! Oh!" "Well—have you read the paper?" "No—just the title." I forget what it was, but it gave the impression that we found something. Look, I spent six months looking at the data—I did a data analysis. Why shouldn't I publish what I found?

Another passage from the Laughing Seafoods conversation makes the same point.

Collins: Let me intervene in this again. What [Frascati1] was saying earlier was that he doesn't want to publish this as a claim for gravity waves, but just publish as—you know—"we found this data." But what you're worried about—tell me if I'm right—is that in fact people won't read the paper this way. People will read the paper as a claim for gravity waves whatever it says.

Unidentified speaker: Sure—people want to reduce it to yes or no.

Collins: People will reduce it to yes or no.

LSU2: That would be my guess.

Collins: OK, but hang on. Let me ask this question: Why are you worried about that? . . . Why are you worried if everybody does read it in this way? You know, if it's not written but they read it, what's the problem?

LSU2: I'm thinking about it. One is, you could certainly expect a lot of ridicule from your colleagues at some point—maybe not overt but covert—if they find the result somehow ridiculous. I mean, you know, everyone values the good opinion of their colleague. . . .

Frascati1: This is true

There are no Americans among the gravitational wave research community who are sympathetic to an open evidential culture, and some of them, perhaps in keeping with a strong individualistic ethos, despise it. The most forceful expression of this point was put to me as follows:

It was mostly at these general relativity meetings. . . . They would give their presentations in such a way that they would lead you, they would show you this data and they would show you the events, and they would show you some statistics they'd done, but never enough of it so that you could really get your arms around it. And they left you with this tantalizing notion that they could go either way. They either could make a claim that they had detected something if they had wanted to, if they had gone the next step in their presentation, or they could back off and say, well, yeah, maybe the statistics isn't good enough. And they left you at that critical juncture. . . . What would happen is that you had to draw the inference. Now that gave them the freedom at any point later on, or maybe even . . . to say well if we choose to say this, we have detected it, or if we choose to interpret that way we haven't detected it because the statistic isn't good enough. It was this . . . ambiguity—OK? That got to me—OK?

Another quotation expresses the point still more strongly in the context of anticipated future problems associated with the development of the World Wide Web.

The problem with unedited data dissemination is that it is irresponsible. The person who was there when the data was taken, who knows the idiosyncratic parts of the apparatus, who has experience with the kind of statistics that comes out of the gadget, and yes, the person who has something to lose if he (she) is wrong needs to validate and then disseminate the data. . . . It has all to do with plain old honesty. When one publishes data and a result, it has a pedigree associated with it, and your reputation as a scientist rests with how well you have analyzed and interpreted it. It should not be possible (as has been the case in this field before) to equivocate and claim validity for an observation at one moment and then deny it later when challenged.²¹

This member of the community believed that the potential ambivalence in an "open" research claim could too easily be used to the researcher's advantage. In this view, a researcher could stake a claim, as it were, to the Nobel Prize, while not being exposed to the risk of being wrong. Less forceful Americans, such as the members of the LSU group, also disagreed strongly with the Frascati Team's approach.

We just don't agree on the theory of data analysis . . . when he says, "I'm not claiming gravity waves, I'm just putting out what I have." And that's perfectly acceptable, except for the fact that if you look at Weber's early papers, . . . "I wasn't claiming gravity waves, I just said what I had," and it is that legacy, I guess, in American gravitational wave physics, that has me very, very cautious of doing that kind of publishing. . . .

And that's never gonna change, because it's just a different way of looking at the physics we do; and we just have to understand that we, as individuals, are different. And that's where your problem comes, actually, as a sociologist, and you're trying to say, "OK, what is it that enters into these things?" And it is much more than just the science, as you know.²²

HISTORICAL AND INSTITUTIONAL CONTEXT

Both Frascati and LSU have been involved in the search for gravitational radiation from the early 1970s. They use "resonant bars"; this, as explained, is basically the same technology as was pioneered by Weber from the late 1960s onward, but the Frascati and LSU bars are cooled to liquid helium temperatures or below. Such devices are known as "cryogenic"

²¹ This comment was made in the context of a discussion about the propriety of broadcasting detector data on the Web. It expressed, then, a general concern about relaxing the distinction between the inside and the outside of the laboratory.

²² Respondents almost invariably think of clashes between scientific cultures as matters of the diverse personalities of individuals.

bars. The LSU group gets all its funding from the U.S. National Science Foundation, and its members are typical, research-active university faculty. The Frascati group is located within the complex of government-supported research facilities at Frascati, near Rome, and draws its members from various universities. The LSU group is the only state-of-the-art resonant bar group left in the United States, while, at the time of writing, the Frascati group was the most advanced resonant bar group in Europe and possibly the world.²³ The group based in Perth—which will figure later in the discussion—also runs an advanced cryogenic bar.

Up to early 1997, the time at which this account is set, the LSU group had never been party to any claim to have seen signals consistent with gravitational waves; over two decades, they had offered only upper limits: "Gravitational waves can be no stronger than such and such, because our bar is of such and such a sensitivity and has seen nothing." The Frascati group, on the other hand, has issued at least two positive reports jointly with Weber in Maryland: one published in 1982 (Ferrari et al.), and one published in 1989, which claimed to see coincidences associated with the rare, "local" supernova that exploded in 1987 (Aglietta et al. 1989). Both papers have been rejected by the wider community, with the supernova paper attracting particularly harsh criticism (Dickson and Schutz 1995).²⁴

The Perth group has recently joined with the Frascati Team to report another positive finding. By February 1997, the Perth group had made at least two conference announcements to the effect that they had recently found strongly suggestive coincidences between their bar and that of the Frascati group.²⁵

Since the late 1980s, the field of gravitational radiation research has been transformed by the advance of a newer, bigger, and much more expensive technology. Research groups at MIT and Caltech combined to lobby for funding to build giant laser-interferometers, which should be more sensitive than the resonant bar technology.²⁶ After a long campaign, funding for the American effort was granted by Congress in 1992. The

²³ This can be said because their most advanced bar was cooled to a lower temperature than any other bar. In 1997, they were joined by a group in Padua using a bar of almost identical design. Weber still keeps two room-temperature bars working. According to Weber's theory, the transducer design of his bars, which differs from the design used by the cryogenic bar groups, means that his devices are more sensitive than theirs. This view is not known to be shared by anyone else.

²⁴ Subsequently strongly rejected by the Frascati group.

²⁵ Remember—the "Frascati group" is the whole laboratory—the "Frascati Team" is the subset that occupies the experimental animus corner of evidential culture space.

²⁶ The resonant bar groups believe that improving their technology would lead to instruments that would be at least as sensitive as laser interferometers within narrow wavebands, even if they cannot match the broadband sensitivity.

American version of the new technology is known as LIGO—which stands for Laser Interferometer Gravitational Observatory—and will cost in excess of \$300 million. The earliest that LIGO might be expected to see gravitational radiation signals is 2002, but most would accept that a few more years are likely to elapse before the devices have a good chance of seeing the waves. The American effort comprises two interferometers, one in Washington State and one in Louisiana (not strongly connected with LSU), each with arms four kilometers long. A European counterpart, with arms three kilometers long, is being built near Pisa by a combined French and Italian collaboration, and there is a German-British group building a 600-meter apparatus near Hannover, while a Japanese group has a 300-meter device. The huge expenditure on interferometry compares with the resonant bar effort, which has probably cost less than \$30 million for the three-decade-long worldwide effort.

From the point of view of those who still experiment with resonant bars, the growth of LIGO has meant that the American environment has become very unreceptive to claims that include any element of speculation. Nowadays, LIGO is the dominant force in American gravitational radiation astronomy, and I argue that it increasingly influences the activities of the bar community in a number of ways.

Reputational Drag

Because of the history of the controversy over Weber's work and the funding battle over LIGO, American researchers are sensitive to insecure discovery claims that might diminish their reputation in the eyes of the wider scientific community. LIGO gained financial support from Congress only after a long and difficult campaign. Even then, in the early years, those working with LIGO did not feel that its future was assured. Under the U.S. system, additional expenditure has to be approved every year, and the demise of the Super-Conducting Super-Collider, after substantial outlay, left researchers feeling vulnerable; a number of physicists migrated to LIGO from the dead Super-Conducting Super-Collider project, with this memory burned deep. Many believed that the difficulty associated with the initial funding of LIGO was partly a result of the controversy over Weber's results. They felt that the future of LIGO could be jeopardized by what they saw as the potential for further reckless reporting of incorrect results by members of the bar community.

Large funding decisions are not made by scientists but by members of Congress and other nonspecialists who may not see the fine distinctions within scientific fields. It is easy to associate contentious results announced by one set of gravitational radiation researchers with that of another set engaged in the search for the same phenomenon, even if they use an en-

tirely different kind of apparatus informed by a different philosophy. Indeed, sometimes physicists themselves seem to have difficulty in making these distinctions. The following interview extract is taken from my conversation with a leading member of the LIGO research effort based at "TECH":

Respondent: Anyway, so what we had to fight was that problem, the problem that this was a field for crazy people; that was number one. And none of the academic places liked it, [TECH] in particular. They looked at this, and they said this was for the birds. Here I was pushing like hell to do this by another technique [interferometry], and I couldn't get any backing in this place, intellectual backing. And the reason was, in part, because it had this very speckled history, and it looked, also, very difficult to do. . . .

You see, I lived in this hostile environment here at [TECH], and they were just gonna' say, "See, you guys are all crazy, and here's more of this bullshit that you guys are presenting to us. How could you do that." . . . I was hoping it would disappear.

Collins: So because you think . . . "Wow, this is amazing" . . . so because, let's say, [named resonant bar researchers] made another one of these crazy claims, let's call it, this endangered LIGO?

Respondent: Yes.

Collins: Just explain it to me again, because I find it very hard to comprehend.

Respondent: Well, we were fighting with the astronomers still at the time, very much so, about the value of doing LIGO in the beginning, and we were continually being tainted by, "You guys are all nuts!" And this just was further evidence of it. So, I mean, they lumped us together with that.

Collins: Yeah, but people surely would have known enough to know that you were separate groups.

Respondent: Doesn't matter. This's a field. "Look—it's pathological science again." That's where it came up again. You see. Now, maybe I overreacted, but I felt that . . . if [named resonant bar researcher group] does something cuckoo . . . we're all gonna get it.

Thus, given that positive reports by the resonant bar community are almost certain to be false according the theory that has informed the design of LIGO, they can only bring trouble or ridicule for a project that is already unpopular among the majority of astronomers. In the early days, this could have damaged LIGO's funding, and it can still damage LIGO's credibility. In contrast, Italian interferometer scientists expressed no such fears, even though they use the same technology.

Big Money Strain

The American gravitational radiation detection community can be thought of as made up of three components: there is Weber's remaining small effort that has been almost completely marginalized both intellectu-

ally and in terms of financial support (it was initially funded by NSF); there is the remaining state-of-the-art resonant bar community, centered at LSU; and there is the LIGO group, centered around Caltech and MIT, which has turned the field into a "big science."

On the face of it, the bar groups should have none of the worries about funding described in the last section; they have always been a relatively inexpensive speculative enterprise. Nevertheless, the LSU group shares the cautious approach of the rest of the American community. Perhaps part of the reason is that they feel they have to be sensitive to the concerns of the interferometer groups because they share a common funding source. Because LIGO is the largest project ever funded by the National Science Foundation, its success and failure represent success and failure for NSF policy. It is almost inevitable, then, that LSU is sensitive to LIGO's fears and what they perceive to be the Gravitational Physics Program Director's interests.²⁷ As one American scientist put it to me:

If there is any one single individual in my opinion in the United States that has influenced the direction and the attitudes that people have on gravity waves, it would probably be [the program director], because he has been the man in charge of funding research.²⁸

and again:

[The program director] does have very definite views, and by being in the position of power that he is and by the way that the American funding system works, why, he has a good bit to say on what work gets done—no way to avoid it.

Members of the LSU group did report to me that they had felt, in the past, that their project would have been in jeopardy should they have reported anything too dubious.²⁹

²⁷ Under the American funding system, decisions are made as much by powerful program directors as by the committees they appoint. There are good and bad aspects to the procedure. On the whole, scientists prefer NSF program directors, who are well-informed members of the relevant scientific community, to committees. Furthermore, individual program directors are more likely to make brave decisions—for example, over interdisciplinary work or adventurous lines of research—than committees ruled by majorities. But program directors' preferences are understood by the community, and when decisions go against a favorite project, disappointment may be personalized while accusations of "bias" are more easily directed at an individual than at a committee.

²⁸ To echo the point made in the previous footnote, this respondent went on to say "and he is a respected theorist himself, or was. One of his calculations was one of the really fundamental ones to show that gravity waves should exist, before he became a science bureaucrat." At the time of writing, this particular program director was taking a research sabbatical, doing more theoretical work.

²⁹ Though, as I will go on to say, this effect, if it exists, is a subtle matter—it is not a matter of crude tailoring of results to fit self-interest.

Collins: You said earlier that there was a stage when you felt that if you had released findings, you felt your grant might be in jeopardy.

LSU1: I know, absolutely, it was, earlier. There wasn't any temptation, because very early on, our experiment didn't work well enough to be happy with it; but had I tried to say that there was anything, we never would have—it was one of those cases where, as [another member of the group] said, it would be so improbable that everyone would say this man cannot be trusted.

At the time of writing, the Louisianians are continuing to seek new funding in an atmosphere of increasing difficulty for resonant bar research. Thus, in so far as U.S. science funding discourages speculative reporting in this area, the pressures have not gone away.³⁰

The structural and institutional pressures on the Italian labs are more ambivalent. The Frascati laboratory seems relatively rich, and its funding seems secure so long as it is seen as *productive*—that is, producing results. The duty to publish that they feel is understood both within and without their laboratory:

LSU1: [The leader of the Frascati lab] argues that he's had people funding his research for all of these years and he really owes them papers.

Frascati2: You cannot decide not to publish the results of these experiments. There are many people working hard, taking data, making the detector work—then you cannot expect that people don't want to publish their results. They should be published, otherwise, why are we doing these experiments?

Furthermore, though there is less money overall in Italy, neither kind of gravitational radiation research seems to be in jeopardy. Thus, a second ultra-low temperature resonant bar detector has just been built in Padua, even while the Americans are considering closing down their one remaining resonant bar group. Also, there is far less evidence of financial pressures on the Italian-French three-kilometer laser-interferometer. While the American building program has, according to a number of physicists, been very much “front-loaded” so as to make sure most of the construction money is spent before Congress has a chance to change its mind, the Pisa installation is proceeding in a leisurely way, driven far less obviously by financial pressure. Thus, the Pisa installation has completed its so-called mid-station without even beginning to build the three-kilometer arms—one of the most expensive parts of the project—seemingly secure in the knowledge that whenever they decide to build the arms, the money will be there to finish the job. There is, likewise, no apparent

³⁰ By the middle of 1997, this funding had been initially refused, leaving the LSU group's resonant bar program without a clear future. By coincidence, they recently reported some very interesting continuous wave results.

competition between the Italian resonant bar groups and the interferometers for a single source of finance; relations between the groups of experimenters pursuing the two kinds of program are cordial, without the tensions that attend the U.S. scene. The setting, then, seems to have features of ordered civil service laboratories rather than the hectic competition between universities, between different communities of scientists, and between different congressional interests that characterizes the American scene.

Theoretical Stress

A "big science" initiative, such as LIGO, must promise success to government and its representatives. LIGO represents a step-function increase in funds for gravitational radiation research and promises a step-function increase in sensitivity; the new machine is meant to be good enough to guarantee not only the detection of gravitational radiation but the founding of a new field of gravitational astronomy.³¹ Given that this is not a matter of gradual development based on existing technology, there is only one source for such a guarantee—theory. Thus, during the course of the search for gravitational radiation from the late 1960s onward, the importance of theory has grown. With the argument over the funding of LIGO, theory has come to dominate experiment, at least in America. It is theory that has justified the spending of hundreds of millions of dollars. For this reason, I treat theory under historical and institutional context.

The most widely accepted theory paints a picture in which, in terms of the emission of "visible" gravitational waves, the sky is almost completely black; under this sky, it is LIGO, which has the only chance of observing and understanding any glimmers of gravitational waves that disturb the gloom, and under this theory, it is virtually impossible for the current resonant bars to see anything except extraordinarily rare events, such as might occur a few times a century.

Theory has also come to dominate nearer to the heart of the LIGO project, because at the levels of sensitivity predicted for the first detections, the only way data will be extracted from noise will be to look for

³¹ A physicist respondent pointed out to me that as far as scientists are concerned, there is an ambiguity about the notion of success: building a sufficiently sensitive gravitational radiation detector and finding no waves would be a scientific success because it would show that the existing theories are wrong. "Negative success," however, does not generate further funding (just as tenured posts generally only go to those who find positive, not negative results). An example of a negative experiment is the search for solar neutrinos—the flux of neutrinos coming from the Sun is far lower than theory would predict. As Pinch (1986) points out, however, the first detectors were justified and funded on the promise of finding a large flux of neutrinos

data signatures that fit precalculated templates (e.g., for in-spiraling binary neutron stars). Again, instead of just looking to see what is there, data will not count unless they match a template. Theory is needed to calculate the templates, and maverick theories must be suppressed if the consensus needed for this kind of experimental philosophy is to be maintained.

Finally, by predicting that the ultimate limit of sensitivity, set by quantum theory, is better for interferometers than for bars, theory has changed the balance of credibility. The technology is so far from reaching the ultimate limit that this is not currently a matter of any practical importance but it has still had a powerful effect on the bars' prospects. Theory, in sum, has aggregated influence in many ways; in this field, theory, once the servant of experiment, has become like a massive star with a gravitational field that affects the shape of everything near it.

Bright sky over Frascati and black sky over Louisiana.—In harmony with their evidential culture and institutional position, the Frascati Team feels that theorists' claims about the structure of the heavens should not dominate experimental work. In 1995, I was told: "If we find something, it is because there is something new. Because nature is kind to us."² And again a year later:

I think that—we are aiming to make a discovery. And no discovery can be made if one has a mind which is within a frame that does not give the freedom to think something different, something new. It seems to me so obvious. Actually, we started this experiment with this idea in mind because I said that already, and I repeat it now—if the gravitational waves that are expected are exactly those which are predicted by the physics, nobody will see gravitational waves from now for the next fifty years.²³ I hope we will both live long enough to find out. So, even if the laser [interferometry] people don't say that, I think everybody hopes that the situation is different.²⁴

The Frascati Team, then, seems less bound by the strong theoretical consensus of American physics. The sky seems brighter over the Frascati

²² This assumes that though LIGO in theory will be sensitive enough to see the waves within a decade, development work will be as drawn out as it has been for the cryogenic bars. It could be, however, that cryogenic bars have had such an extraordinarily painful development phase because any adjustment, however small, takes three to four months while the apparatus is warmed and cooled again. Barring accident to the main vacuum pipes, interferometers have an advantage in this respect.

²³ Sam Finn pointed out to me that it is possible to have an adventurous view of the cosmos and still be conservative in terms of evidential practice; it certainly is true that one does not follow strictly from the other, though without an adventurous view of the cosmos, one is less likely to want to release results early.

Team than over most of the United States (Maryland is an exception).³⁴ For example, one member speculated that relatively high energies—enough to provide a half-dozen or so pulses a year—might be emitted by coalescing MACHOs (MAssive, Condensed, Halo Objects). Such objects, which would be invisible because they do not emit heat, light, or shorter wavelengths, might make up part of the “dark matter” component of the universe. These massive but invisible stars could fill the area around our galaxy.

In contrast, there was no such speculation in public among the Louisiana team with respect to the bars. They seemed happy to accept the calculations of the major theory groups that stand behind LIGO. The axis of Louisiana’s arguments turned on the narrowband sensitivity of resonant bars, which would enable them to compete in terms of the observation of certain kinds of sources, the improved sensitivity of further generations of resonant spheres, and the much greater directional sensitivity of resonant technology. But this argument was conducted almost entirely within the framework of the levels of energy and the cosmic scenarios informing the development of LIGO.

Academic Drift

Because LIGO is such a large project, and because the future of resonant bar work is not clear, interferometry groups may represent the best employment market for academic apprentices trained in resonant bar experimentation. One young physicist explained to me his attraction to LIGO in the following way:

I’m certainly just following my gut. I’m certainly following the interferometers. That’s where the money is, that’s where the people are, that’s where the scientific opportunity is, so that’s where I concentrate most of my attention. It’s just that—it’s one of those unconscious choices that—it’s a value judgment. It’s the kind of value judgment that people make when, you know, you decide whether to have children or not, or things like that. They often don’t make those judgments very explicitly, but they know they’re based on some kind of understanding of things.

Given that in the United States it is taken to be the duty of a project leader to find secure employment for graduate students and postdoctoral fellows who move on, it is important for the resonant bar groups to maintain good relations with the rest of the physics community; they must not marginalize themselves intellectually or in terms of the informal academic networks that lead to jobs.

³⁴ Joseph Weber is at the University of Maryland

In the case of the LSU group, an added complication is that one of the LIGO sites is just 45 minutes' drive from LSU, which could bring benefits if relations remain good:

It is our hope that since we're already active in the gravity business, that if they come in and build a LIGO detector here, that this is going to enhance the program here, and it will. We've said that we would make some faculty appointments in gravitational physics and so on. . . . We see our position being with new faculty hires, being someone who would be actively involved in using the detector and also quite naturally that people running it—it's about thirty miles from here . . . there's nothing over there—absolutely nothing. So you're not going to find your Ph.D.-level people living in the country town of Satsuma—they'll base themselves around here. We've got good computing facilities, I think it's just going to be natural.

Furthermore, one of the members of the LSU group is actively seeking to develop research under the aegis of the LIGO project.

The members of the Frascati Team seem well insulated from such pressures, making optimistic plans for further developments in resonant mass technology (resonant spheres are the latest idea). On the other hand, members of the Frascati group who share the mathematical animus, many of them junior, may be more influenced by job prospects and professional networks. The three-kilometer interferometer group may eventually be a source of employment for them, just because of its great size as compared to the resonant bar program. Furthermore, partly because all searches for gravitational bursts are coincidence experiments, the whole community is becoming more and more globalized, and the cultural influence of the U.S. community is bound to be felt by physicists who frequently hold temporary fellowships away from their own laboratories and who interact with all the members of the gravitational research community several times a year at conferences—interacting, for purely demographic reasons, with far more interferometer scientists than resonant bar scientists. Globalization is, of course, well advanced among high-energy physicists who now form a sizable proportion of gravitational wave researchers on both sides of the Atlantic.

The Growing Influence of High-Energy Physics

American physics practice is led by the elite, competitive science of high-energy physics. Its influence on gravitational wave research has become marked as LIGO's personnel have become dominated by high-energy physicists.

From the early days, Weber's data analysis practices have been unfavorably compared with those of high-energy physics. In my interviews, I find American groups holding up high-energy physics as the reference

group for how to do good data analysis. Members of the Frascati group, on the other hand, were less impressed, believing that the delay histogram method (see below) was better.

No. The concept of standard deviation is used by most physicists, particularly the high-energy physicists. It is not the concept which I think is appropriate for our experiment. . . .

But if I convert this to standard deviations, and say four standard deviations, then any high-energy physicist will say "not enough," because they think in a different way.

Here we see the clash between choices of evidential threshold in high-energy physics and resonant bar work.

Sharon Traweek's (1988) anthropological study of high-energy physicists reveals some of the features of that community that encourage evidential individualism. For example, she repeats a story of a physicist who prematurely reported the discovery of an "exotic meson" before it had been confirmed with sufficient statistical rigor. The man angered the community and lost a promotion (p. 118). She also reports that high-energy physicists say that even if one has absolute trust in an experimentalist, one should multiply his claimed amount of possible error by at least three (p. 117).

Traweek also reports that American high-energy physicists from different groups do not talk to each other very much, while interaction is needed to reduce distrust of the methods of the resonant bar community, such as the use of delay histograms. Members of resonant bar groups have complained to me about the initial arrogance, competitiveness, and unwillingness to communicate of the new entrants from high-energy physics (a problem that seems to be fast resolving itself).

Without being evaluative, it does seem that there is a difference in approach between high-energy physicists and certain resonant bar groups. It has to do, I would suggest, with the relative "normalcy" of high-energy physics as a science. Over many years, as a result of experimental bravado combined with engineering excellence, new particle after new particle has been produced. A certain pattern of expectations for the nature and duration of the experimental cycle has been built up, and with it, a well-defined image of what constitutes a mistake arising out of an insufficiently rigorous approach. The relationship between theory and experiment has been fruitful, with the experimentalists showing time after time that, given the resources, they can build the machines that will see anything that theorists say it is possible to see. In contrast, the whole history of gravitational wave research has been a matter of looking for something that ought not be there. All the ingenuity has been directed at struggling to keep going while trying to do the impossible rather than regularly doing the possible. Any signals have to be extracted from noise using any and every means;

the problem has been not to drive away false signals but to maintain some grounds for hope. This has given rise to a very different experimental culture. The pre-high-energy physics culture has survived in Frascati, while the Louisianians predilections have been reinforced by the influence from high-energy physics.

INVOLUNTARY BLINDING AND THE PERTH-ROME COINCIDENCES

We have explored the cultural and institutional setting of the field of gravitational wave detection. We can now understand the unfolding of two interesting recent events. But first we need to explain some features of data analysis.

Analysis of Coincidences

Delay histograms.—The trouble with gravitational waves is that they cannot be turned off. It is difficult to do a detection experiment on a phenomenon that cannot be turned off, because it is hard to compare a control situation and an experimental situation. Ideally, one would like to be able to shield one's apparatus from gravitational waves from time to time. Then one would compare the output of the detector in its shielded and unshielded periods; any excess of activity in the unshielded periods compared to the shielded periods would be a strong indicator of the existence of gravitational waves. But since gravitational waves are distortions of the very fabric of space time, which lose only an infinitesimal proportion of their energy to matter, they penetrate all imaginable shields.

The invention of the "delay histogram" represents a partial solution to this problem.³⁵ The output of a resonant bar detector is a record of a series of levels of energy at small time intervals over a period of days, weeks, months, or years, depending on how long the detector remains working reliably. Other groups can take this data stream and compare it with the output of their own detectors, looking for coincidences of levels of energy or, more usually, coincident, sudden changes in levels of energy. The way this is done is to compare the two data streams at a variety of temporal offsets, or "delays." If the two data streams being compared were recorded, say, one week apart, then there can be no coincident changes in energy caused by outside sources—any apparent coincidences must be a result of chance combinations of the inevitable peaks of "noise" of the two sys-

³⁵ This interpretation of the delay histogram protocol—in terms of its being a substitute for shielding—is my own invention

tems. To a first approximation, the same goes for every time delay—positive or negative—except zero.

Actually, because the two detectors are typically separated by several thousand miles, a signal received on one may not be “seen” by the other until several 100ths of a second have elapsed—the time taken for signals, which travel at the speed of light, to traverse the distance between them. More important is that the response of resonant detectors to inputs of energy is “smeared” over time—they experience an energy “kick,” which may take time to register and time to die down. Thus, the definition of coincidence is another point of debate and negotiation. But, if we take “zero-time-delay” to mean “within a second or so,” we will be able to understand the techniques and the arguments.¹⁶

Now imagine that we count the number of coincidences between the outputs of two detectors seen over a period such as a month. Then we take the recorded data streams and offset their starting points by, say, one minute and count again; this time we will be counting apparent “coincidences” rather than real coincidences. We then repeat the process with an offset in the other direction—minus one minute instead of plus one minute. We can continue to try as many positive and negative “delays” as we like. The array of numbers produced by this process is plotted on a “delay histogram.” The height of the “bins” represents numbers of “coincidences,” while the horizontal scale represents delays, negative and positive. The zero-delay bin is in the middle. The heights of all the non-zero-delay bins should fluctuate randomly around a mean level that represents the average number of apparent coincidences that might arise from chance matchings of peaks of noise. If, however, there are some peaks of energy that are due to outside signals impacting on the two detectors simultaneously, as in the case of gravitational waves, we would expect these to show up only when there was a zero time-offset. Thus, the central, zero-delay bin in the delay histogram should be higher than the other bins if genuine signals are present. A coincidence experiment compares two data streams in this way and looks for a “coincidence excess” in the zero-

¹⁶ Apart from the physics of the meaning of coincidence that I have just described, simple, accurate timekeeping seemed to be a source of trouble over and over again in this field. One would like to be able to compare the time at which events occurred to within an accuracy of a 100th of a second or so, but at some sites, clocks were used that could drift by a second or two per day before they were readjusted. One group, so it was reported to me, used their computer clocks, correcting them a few times a day by telephoning a local radio station. When I expressed surprise at the problems encountered with timekeeping, a respondent explained: “Everyone was working hard just to make their experiment work, and there was little thought given, in the early years, of just how coincidences would be done. Then, when it was necessary to do coincidence work, it was hard to force good timekeeping into the historical way that each experiment’s data taking had developed.”

delay bin. The height of the zero-delay excess shows how many gravitational wave events (or other events that mimic gravitational waves) have been experienced by the bars (see fig. 2).³⁷

Once we think of a gravitational wave experiment as comprising two detectors rather than one and we think of the effect of gravitational waves as being coincident excitations of two detectors rather than excitations of a single detector, we have a way of turning off the effects of gravitational radiation on the experimental apparatus. The effects are turned off in all the non-zero-delay bins and turned on in the zero-delay bin. Unfortunately, this is not a complete substitute for true shielding, because it is not really gravitational radiation signals that are being turned on and off but the influence of anything that can effect both bars simultaneously.³⁸ Nevertheless, it is this invention that makes the experiment, dealing as it does with tiny effects, possible.

The delay histogram method of analysis is very seductive. Even a small zero-delay excess, when set among a large array of delayed bins, looks convincing. The experimenter is bound to ask, if there is nothing there, why did the data-analysis algorithm choose this particular central bin to be the highest one? A quotation from an early interview (1975) is revealing:

More recently, two Japanese scientists have repeated the experiment, again using the amplitude algorithm. They did see the largest number of coincidences in the zero-delay bin, with a factor of between one and one-and-one-half standard deviations.

By the standards of modern physics, that's not a significant effect. Nonetheless, the fact that their computer discovered the largest number of counts at zero-delay may be significant

Statistical massage.—It is a well-known psychological phenomenon that experimenters tend to analyze experimental data in such a way as to favor their own hypotheses. Nearly always, this is not a matter of dishonesty but unconscious bias (e.g., Rosenthal 1978). Given the nature of gravitational wave detection experiments, there is ample scope for inadvertent bias in the analysis of data. Any delicate apparatus will have noisy, insen-

³⁷ It seems that Joe Weber was the first to use the delay histogram technique. Guido Pizzella told me that Joe Weber invented the technique while Pizzella himself gave it the name of "experimental probability."

³⁸ To give an idea of other effects that have been considered, it has been suggested that two bars separated by the width of the United States might be influenced by the simultaneous switching on of many electric pumps across the United States in response to coordinated toilet flushes during the commercial breaks in televised *Monday Night Football*. I believe I heard the gravitational effects of the simultaneous movement of the massive amounts of water also being discussed in relationship to the interferometers.

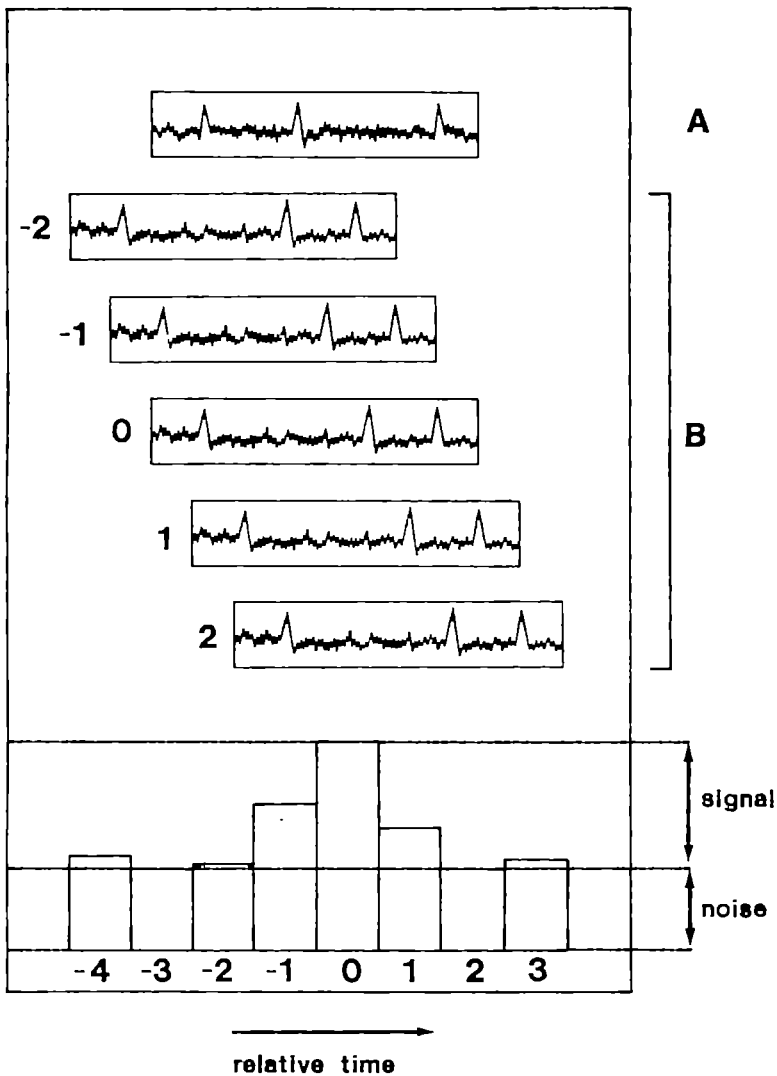


FIG. 2.—Signal extracted from noise by delay histogram

sitive periods and quiet, sensitive periods: the outside world of cosmic ray showers, tides, and earthquakes may vary in its intrusiveness; the immediate environment, both human and physical, may affect the apparatus in many ways; and the internal components, such as seals, electrical connections, liquid gases, and the metal itself, may harden, crack, leak, cease to insulate, and stretch and strain. The tiniest change in any of these

things affects the immensely delicate detector. Every experimenter has to decide when the apparatus is working satisfactorily, so that data are a candidate for analysis, and when it is not, so that the data should be rejected at the outset (Holton 1978). Other parameters that can be changed are the length of the period that is taken into account as a bin, the nature of the initial signal-processing filter, and where the threshold for what counts as a potential signal is set.

The following extract is from my interview with a member of the LSU group in December 1996:

If our apparatus is not working correctly, as we can tell—the noise temperature going up—we'll do a new calibration and build a new filter, and then on the basis of that filter, we would begin recording events again. . . .

Each filter would be different for a different apparatus. And this is something that [named groups] have not understood. [Named group] have felt that they could build a filter on a theoretical model and that would be good; and in theory, you could do it, but in actual fact, you don't understand the actual operation of the apparatus well enough. So it is more complex than one would initially think.

Even this most cautious of experimenters recognized the need for continual monitoring and readjustment of the apparatus in light of its performance. Yet retrospective readjustment of experimental parameters always offers the possibility for the artificial construction of a nonexistent signal—that is, “statistical massage.”³⁹

Given the history of the field, American groups are very sensitive to the possibility of accidental statistical massage. One of the events that caused Weber to lose credibility was his reporting of coincident signals between his own antenna and that of a group at the University of Rochester. It turned out that because of a mistake in timing, the two sets of signals had starting points that were more than four hours different, so Weber had accidentally constructed a “signal” out of what should have been pure noise.

Most scientists, and certainly almost all American and British scientists, believe that the Frascati group's report of signals coincident with Weber's apparatus at the time of the 1987 supernova was a result of unconscious post hoc adjustment and readjustment of parameters until a spurious “signal” emerged. This has not helped the Frascati Team's case in spite of their strong rebuttal of the charge.

³⁹ It is tempting to think that, following standard psychological practice, once parameters have been set in “preliminary runs,” they could be frozen for “experimental runs.” Indeed, this was a justification offered for publishing with low evidential significance (see below)—so that the parameters recorded in the paper could be publicly certified before “freezing.” It seems, however, that the experiment is likely to be too delicate to allow “freezing” in the foreseeable future.

Thresholds.—The scope for statistical massage increases with the number of data points to be analyzed, and this is affected by the threshold that is set for potential gravitational wave events. Two (or more) streams of data are to be compared. Each of these streams may be visualized as a squiggly line with almost random fluctuations up and down; there will be big fluctuations and small fluctuations. Let us say that the inputs in which one is interested are sudden inputs of energy, represented by sudden upturns in the wavy line. The question is: How big must an upturn be in order to count as a potential gravity wave signal? A threshold must be set for the events that are to be compared in the search for coincidences. If the threshold is high, only a few events will be compared. As the threshold is lowered, more and more candidate events will be compared for temporal coincidence.

As might be expected from the Laughing Seafoods conversation, on this point too, the Louisianians take a view consonant with their closed culture, and the Frascati Team takes an open view. The LSU group argued that the threshold should be set so high that very few events were being compared; this would mean that even a single excess event in the zero-delay box would stand out. They argued that, in theory, expecting to see even one event a year on antennae with their sensitivity was extremely optimistic and that the threshold should be set accordingly. Furthermore, they argued, the lower the threshold and the more candidate events being compared, the greater was the scope for any apparent zero-delay excess to be the outcome of unconscious bias in the data analysis; the more data there were to be analyzed, the more opportunity was there to “dig down” and pull something out of the data that was not really there.

The Frascati Team argues that finding one event would be meaningless anyway, because a single event could always be the result of chance and could have no statistical credibility. Such an observation would, therefore, require replication, and if gravitational events are as rare as they are expected to be for instruments of that sensitivity, decades might pass before such a thing would appear again. The high threshold, low event frequency search program is a recipe for finding nothing that can be believed until long after the interferometry program has been completed. In any case, if the threshold is set high, in theory, the amount of energy required to produce the single event would be so great that astrophysicists would refuse to accept it for that reason.

As we have seen, the Frascati Team argues that if the resonant bar program is to have any chance of success, it must be based on the assumption that there is something unexpected in the heavens. What follows is that the threshold should be set low enough for any zero-delay excess to imply a frequency of events, such that repeated observations can be made within a reasonable fraction of an experimenter’s lifetime. In other words,

most of the data should be treated as potential signal and exposed to scrutiny; it should not be treated as noise and hidden behind a high threshold.

Involuntary Blinding: The Louisiana Protocol

We are now in a good position to examine one of the most interesting episodes in the history of the search for gravitational radiation. As explained at the outset, coincidence experiments force groups with different cultural propensities to work together.⁴⁰ In spite of the contrasting institutional, financial, and cultural backgrounds of the Frascati Team and the Louisiana group, they must cooperate to produce results. How can Louisiana maintain its preferred strategy when the Frascati Team could take Louisiana's data stream, compare it with their own under their preferred analytic protocols, and in a spirit of evidential collectivism, issue high-risk reports with low evidential significance?

In the early 1990s, the LSU group introduced an intriguing new data exchange protocol to try to get around the problem. From that date, they decided to release data from their own experiment only in a disguised form. The Louisianians' innovation was as follows: Initially, in 1991, they began to release two data streams, one genuine and one false, without saying which was which. Later, they began to release their data streams in the form of a continuous loop with 1,000 "starting points" marked on it. Of these 1,000, only one was genuine, while the other 999 were false. They would not say which was the true starting point. Any other groups receiving their data had to find out for themselves which point represented the start of the data stream.

To do this, other groups would have to construct a delay histogram with 1,000 bins corresponding to the 1,000 potential starting points on the Louisiana data loop. Of these bins, 999 would be the result of comparing data trains with a time offset, and only one would exhibit zero-delay. If they discovered that one bin out of 1,000 was higher than the rest, they might choose to claim that they had discovered a zero-delay excess—that is, a genuine signal. But they would not know if this bin was really the zero-delay bin or if it was another bin with an excess due purely to coincidences in the noise. To know whether they were right, they would have to depend on the Louisiana group. If the second group were to announce a "result" without consulting Louisiana to find out if they had found the right starting point, they would risk making public fools of themselves.

⁴⁰ Note that the principal members of the LSU group and the Frascati group have been friends for many years not only professionally but at the level of personal visits between families; the drama was played out with a minimum of personal animosity.

In effect, the Louisiana data exchange protocol forces any group who wishes to cooperate with them in a coincidence search to run their experiment "blind," whether they like it or not.

Initially, the invention of this protocol, with only two data sets in the first instance, was on the basis of a mutual agreement between LSU and Frascati. It also appears that the Frascati group was initially happy with the extension to 1,000 starting points, since any result based on this protocol would look sound. More recently, however, the Frascati Team has become much less happy. They consider that to continue with the protocol serves the interests of the interferometer groups rather than the bar groups. A member of the Frascati group put it this way, overlooking, perhaps, their initial ready agreement to the protocol:

If we write a paper on coincidences, we might damage the LIGO project. So [the NSF gravitational physics program director] asked the Louisiana people to be very careful in writing these papers. So we have a lot of difficulties in exchanging the data with Louisiana. And finally we reach this agreement: . . . So when searching coincidences—if we find coincidences with these wrong files, it certainly is not good. This is a very clever point.

In fact, the NSF program director did give strong support to the new protocol. As he put it:

I had talks with the people at Louisiana before they got involved with data exchanges. I said, "You've gotta protect yourself, you don't want to risk errors." And they knew that too. They were asking, "How much garbage shall we put in with the good stuff? Should we put in a factor of five or 10 or 100?" And I was ultra-conservative, and I suggested massive amounts of nonsense. And I think they got a practical compromise.

In a later interview, he went on to amplify as follows:

Collins: Who invented that LSU protocol?

Program Director: Well [a member of the LSU group] was talking about this, and I remember having a discussion with him where I asked him what would the chance be if you were exchanging your records blind. . . . "What's the chance that you can get a false alarm?" And he came up with a number which is a few percent. And I said that he should be very careful, and that if I were in his shoes, I would certainly not want it to be more than one in 10,000 or 100,000. And he said, "Well, you have this problem, you have to do data analysis—massive amounts of data analysis, and it gets to be a headache." And I said, "Who cares [laughter], we're playing for big stakes, and we don't want false alarms." But I think he was getting similar advice from lots of people. . . . He worked out some procedure.

The advantages of involuntary blinding.—As far as physicists inclined toward a closed evidential culture are concerned, the major benefit of

compulsory blinding is that no chaos can emerge from irresponsible interpretation of the Louisiana data. Anyone can use it (in principle), but it would be a foolish person who broadcast a result of a coincidence analysis with the LSU data without checking first that what they believed to be the zero-delay box was indeed the zero-delay box. Thus, LSU is effectively enforcing its own evidential culture—compelling the struggling convicts to run in their chosen direction.

Of course, if a blinded analyst does find the right box, this is powerful confirmation that the effect is real and not a result of post hoc statistical massage. Any post hoc tampering to increase the height of a favored bin may be enhancing the wrong one. The procedure seems, then, an ideal way of eliminating both “crazies” and unconscious biases.⁴¹ A third party, a partner in the laser interferometry effort, put it this way:

If he's doing his statistics right and claiming a coincidence at a believable level, he'll have no problem picking out which of the thousand is the right beginning time. But if he's fudging things—even unconsciously—then they'll fudge on the wrong one, or something, or won't even be motivated to start fudging because it will get lost in a morass. I think it's very clever. You know, not all social or psychological problems have technological solutions, but I think in this case, someone was clever enough to have come up with a technical solution for a psychological problem. So I think it's beautiful. . . .

Scientists from other fields would say, “Oh look, those gravity wave people can never do anything right, here's another piece of crap from them. . . .” Claims that are believable—whether they're true or not of course is very subtle as you guys [sociologists] know—but they have to be believable. So this is a way of getting around the biggest hurdle of believability, namely, the incredible temptation . . . of massaging the statistics. This will make it virtually impossible, and furthermore, it makes it certifiably virtually impossible for anybody from outside who would look at it; that's why I think it's so clever. So that either there won't be a false find or there might be a believable claim—either way, it's better than being almost certain of an unbelievable claim.

It is true that under this protocol the Frascati group has made a few wrong guesses about the starting point. They have, however, checked these out privately with the Louisiana group before publishing. From the point of view of anyone not working in gravitational waves, it looks, then, as though the protocol is excellent; it looks especially good to anyone working on interferometers. As one scientist put it to me, stressing the view of the interferometer community:

Respondent: And they've had several times when [a member of the Fras-

⁴¹ “Crazies” include complete outsiders who might take the data from any pair of laboratories, perhaps from the Internet, and analyze them in their own preferred way, without knowing anything about the experimental apparatus

cati group] has said, "number 62—that's the one," and they [Louisiana] said, "Nope!"

Collins: That's happened a few times has it?

Respondent: Yep! . . . "Oh well it wasn't 62, it was 54" . . . "Now wait a minute . . . how many guesses are you going to get before we invalidate the procedure?"

The disadvantages of involuntary blinding.—The Frascati Team confirmed that they had indeed made some wrong guesses. Is involuntary blinding, then, simply a brilliant solution to a "psychological" (that is, sociological) problem? Not according to the Frascati Team; for them, cutting risk to near zero is not the best policy. Going back to the aim set out at the beginning of the paper, we can explore their rationale.

First, the 1,000 starting point procedure imposes a lot of work on the recipient of the data, because he or she must construct a 1,000-box histogram—in the normal way, a delay histogram would contain many fewer boxes than this. More serious, a member of the Frascati Team pointed out that the only effect that an analyst can discover under this protocol is one that is strong enough to have occurred by chance less than one time in one thousand (i.e., more than three standard deviations above background). If the effect is less strong than this, one or more of the non-zero-delay bins is likely to show a higher peak than the zero-delay bin, purely as a result of chance. For example, if there is a genuine effect, such that there is a 1% chance of its being due to chance (2.5 SDs—a result that would be considered more than adequate in the social sciences), the 1,000 starting point protocol is likely to disguise it among around 10 results of equal or greater apparent significance. Thus, work on a 1% signal is impossible. In physics, 2.5 standard deviations is not much and would not normally be thought to amount to a significant result, but this is not a normal field of physics; we know we are going to stay near or within the noise for a long time to come, and therefore, we know it will be a long time before we escape from the developmental phase of the experiment.

A member of the Frascati Team argued that it is necessary to continue to understand the phenomenon and develop the apparatus in appropriate directions, and for this purpose, it is necessary to work with any signals that might be there, however poor their statistical significance: "I still think it's important that we find 1% excess, that you cannot find it in that way because you require better than one per 1,000. That's the problem." To enhance the signal, in a situation where the signal and the apparatus are so ill-understood that it is impossible to design the best filters and set other experimental parameters to their optimum level by prior theoretical considerations, one should adjust the parameters in such a way as to enhance the height of the zero-delay bin, and to do this, one has to know which it is.

Interestingly, this argument carries forward two positions found in the early days of gravitational radiation research. Joe Weber, who was trained as a naval radar operator during World War II, first organized his data analysis on the basis of a radar search. Here one continually tunes and retunes the "receiver" and the analysis parameters until the signal is enhanced as far as possible using the signal itself as a reference point. In circumstances where one is certain that the signal must be there, or that the major risk is not finding a signal that might be present, this procedure is optimal. In circumstances where the existence of a signal is uncertain and the major risk is falsely reporting something that is not there, the procedure is dangerous because it encourages post hoc manipulation.

Thinking along these lines, it is tempting to reconstruct both the history and current situation in gravitational radiation research as a battle between those for whom the major risk is finding a false signal and those for whom it is missing a real signal. The interferometer and bar groups occupy the two poles of the dichotomy. Since the interferometers will not be "on air" for several years, the potential damage caused to them by a false bar signal is their principal worry. The bars, on the other hand, are losing credibility under the interferometers' "strong theory regime" and will soon be overtaken in sensitivity; their very survival could depend on them finding signals fast. Of course, under this model alone, the Louisianians ought to want to run in the same directions as the Frascatians.

In any case, according to the Frascati Team, there are yet more disadvantages to blind analysis. Once one has a signal to work with, one might be able to enhance its value as evidence in ways other than simply boosting the signal-to-noise ratio. One might find, for example, that not only does one have coincidences—albeit at a low level of statistical significance—but the ratios of the levels of energy of the coincident signals are consistent between the two detectors; this would be further evidence for a common source. Or one might find that the signal peak, low though it is in absolute terms, correlates with the times when the antenna is pointed in some particular direction—putatively, a strong source of gravitational waves. These are ways of boosting the credibility of a signal but only after one has a vestigial signal to work with; it can be done well only if the vestigial signal is in the public domain. From the point of view of evidential collectivism, compulsory blinding is a mechanism for suppressing the development of instruments and data analysis techniques.

The Perth-Rome Coincidences

The tensions and differences of view we have described have recently found a new focus as a result of the discovery of the apparent coincidences between the output of the Frascati antenna and that belonging to the

group at Perth.⁴² The new signals, like all those that have been announced, imply an energy level that requires what the scientists refer to as “new physics” or “new cosmology” to make sense; in short, they are as difficult to accept as all the previous claims. Nevertheless, at a conference in Pisa, held in March 1996, the leader of the Perth group suggested that they be announced; the leader of the Frascati Team, who was organizing the session, was less enthusiastic.

Perth: Because we're not down to be giving any talks . . . in that bar session, and . . . I don't want to give the impression that we're not working on the bars. . . .

. . . And so I thought it might be an opportunity just to say that we're seeing some interesting coincidences—we're trying to understand it.

Frascati: If they ask me, I will make no comment . . . Because you see, we are in this difficult situation, since we are publishing papers on the supernova. . . .

Perth: I think it's not right . . . because of the fuss that there's been in the past, to deny the data.

Frascati felt too bruised by the reception of the supernova claims to want to be controversial again, but Perth went ahead and offered the results to a skeptical audience. The leader of the Perth group explained to me much later: “We were not going to be bullied by people who have their own agenda. We believed that what we had seen was reasonable and interesting and that you should tell the story as the story goes—as it unfolds.”

Over subsequent months, the evidence for coincidences between Perth and Frascati grew stronger. Once more, the Frascati Team felt it was time to publish something positive. They believed, however, that, as well as those of the Perth group, the results of the LSU group should be included in any report. That is the context of the Laughing Seafoods conversation.

We have now understood enough of the technical, institutional, national, and cultural environments of these groups to understand the state of the argument in December 1996 when the Laughing Seafoods conversation took place. The immediately pressing point of the conversation and of various other discussions that took place at the same data analysis meeting was whether the LSU group would agree to their names and their data being included in a joint publication that would include a report of the coincident signals between Frascati and Perth.⁴³ The Frascati group

⁴² In fact, the Frascati antenna discussed here is located at CERN in Geneva, but I will stick to this usage to save confusion.

⁴³ The conference was the first Gravitational Wave Data Analysis Conference, held in Boston, on December 6–8, 1996. LSU, it should be stressed, was not suppressing positive data that would support the Rome-Perth coincidences. But it was generally agreed that even the presentation of neutral data by the LSU group in the same paper

also wanted the Louisiana protocol relaxed for the sake of future collaboration and development, and they wanted the threshold for potential signals to be set low. The leader of the LSU group was trying to make up his mind what to do. His position was becoming less clear cut, in spite of the pressures toward the retention of uninterpreted data discussed above. The LSU group had been running detectors for two decades with very little to show for it. For some time, they had been seeking funds from NSF for a new generation of more sensitive spherical resonant detectors. If this development does not go forward, they may conclude, that they should be saying more about what they have done (especially as the leader of the group approaches retirement).⁴⁴ As one member of the group put it in discussion with me:

LSU1: And it may be that [we have] been too cautious, er, in saying we want this to be a real field. And we want to develop the science—the science is important stuff to do because no one really understands gravity yet; they don't understand why it's the way it is, so since we can begin to do these experiments, we have to do them, we have to do them responsibly. And I can't say that I looked at it this way when I started, but on the other hand, on thinking about it, I'm pretty sure that foremost in my mind was there is no way that I want to do anything that is going to have me labeled as a crackpot. . . .

When someone then begins to find experimental things that may or may not be there, then, again, you've got to be very careful, and I suspect that underlying all of this is the desire to really say, "Well, OK, if my work is going to be significant, I've got to make sure that no one thinks that it's crazy."

Collins: Yeah, you've got a gamble to make though, haven't you? Because you can always make absolutely sure that nobody thinks you're a crackpot by never finding anything.

LSU1: Oh—that's true—absolutely—and Weber accuses me of that. In years past, Weber has said, "There are two mistakes you can make—one is to find things—anything—you know, to find coincidences or something like that given any data," and I've said this at a number of talks I have given—you can make the other error and that is to try too hard, and that is what Weber accused me of; he said, "You can also try too hard to not see events." And I've been very aware of that.

Later, discussing whether LSU would agree to publish with Frascati, the conversation went on:

as the positive reports by the Rome-Perth collaboration would boost the credibility of the whole exercise. As it was, Frascati was becoming marginalized.

⁴⁴ An NSF review committee, which met at the end of 1996, decided not to fund the more sensitive detector and to close down all work on resonant bars in the United States once foreign resonant detectors had reached an unambiguously higher level of sensitivity. At the time of the interviews reported here, this was unknown to the LSU group.

Collins: So [to add your name] will add more weight to the paper, but you must also have a worry that if the paper is dismissed, its gonna take credibility away from you?

LSU1: Yeah, but we can stand the shot now; that we can stand. But on the other hand, maybe there is something there—I am very aware of that. And so do you throw away 25 years of working on this? And this is not thinking of a Nobel Prize in this . . . it's just saying, "OK, well we've not published much"—do you want your 25 years of work just to quietly disappear—while other people go ahead and get on with advancing the field.

Now we know what was going on in Laughing Seafoods.

SUMMARY AND CONCLUSION

I have examined the clash between two evidential cultures. The closed evidential culture of the American group of gravitational radiation researchers is homologous with the institutional and structural forces in which they are embedded. The open evidential culture of the Frascati Team fits with the looser set of constraints under which they work, though the growing cultural homogeneity of big science makes things more ambivalent in that case and gives rise to disagreements within the same institutional setting.

I have looked at two recent episodes of resonant bar research in light of these differing backgrounds and cultures—involuntary blinding and the putative coincidences of the Rome and Perth groups. In both cases, the dilemmas that the groups of scientists face about what to count as data and what to count as noise and how to present and analyze their potential findings are set within a network of considerations that stretches in many directions. The decisions that have to be made rest, inevitably, on the sort of reasoning that is familiar in any walk of life. The young physicist quoted earlier who had decided to invest his effort into interferometry made the point in respect of career choice: "It's the kind of value judgment that people make when, you know, you decide whether to have children or not, or things like that." I have tried to show that this kind of tacit decision making, informed by background assumptions, many of which are not recognizable as "scientific," goes into demarcating interesting data from uninteresting noise. The very notion of "data" depends on different scientific traditions and, in this case, patterns of institutional forces.

My guess would be that, in the immediate future, the growing global power of interferometry will result in the further marginalization of evidential collectivism and open evidential cultures; in this field, the one right way to do science will become more and more firmly entrenched. But it will be intriguing to watch events in the first years of the new millenium

as the interferometers belonging to different national teams see their first signals at the margin of noise. For closed evidential cultures, the prospect is not without its problems, and even now, some scientists are trying to organize the way the findings will be made public.

VISIBLE AND INVISIBLE FORCES

I have described the relationship between the evidential cultures of the two groups, their view of the heavens, and their institutional positions as a homology. I choose that word because it would be quite wrong to draw crude causal connections between the structures and pressures within the respective environments and the scientific choices made by the individuals. The history of scientific heroism is the history of scientists resisting institutional pressures, and there is no reason to suppose that the principal actors who speak from these pages are not just as capable of resisting external pressure as a their famous forebears. Of course, given that every scientist discussed here would readily concede that, in the last resort, blinding is the only sure way to eliminate unconscious bias in data analysis, they would also have to accept the possibility that structural forces *might* affect their larger judgments in subtle and invisible ways. It would be very difficult to turn this "might" into anything definite, as the complex workings of law courts reveal. Fortunately, we do not have to prove the matter one way or another for any individual scientist; we have only to describe the pressures and the plausible patterns of action within the scientific community.

REFERENCES

- Aglietta, M., G. Badlon, G. Bologna, C. Castagnoli, A. Castellina, W. Fulgone, P. Galeotti, O. Saavedra, G. Trinchero, S. Vernetto, E. Amaldi, C. Cosmelli, S. Frasca, G. V. Pallottino, G. Pizzella, P. Rapagnani, F. Ricci, M. Bassan, E. Coccia, I. Modena, P. Bonifazi, M. G. Castellano, V. L. Daykin, A. S. Malguin, V. G. Ryassny, O. G. Ryazhskaya, V. F. Yakushev, G. T. Zatsepin, D. Gretz, J. Weber, and G. Wilmot. 1989. "Analysis of the Data Recorded by the Mont Blanc Neutrino Detector and by the Maryland and Rome Gravitational-Wave Detectors during SN 1987A." *Il Nuovo Cimento C* 12 (1): 75-103.
- Ashmore, Malcolm. 1989. *The Reflexive Thesis: Wrighting Sociology of Scientific Knowledge*. Chicago: University of Chicago Press.
- Bloor, D. 1973. "Wittgenstein and Mannheim on the Sociology of Mathematics." *Studies in the History and Philosophy of Science* 4:173-91.
- . 1976. *Knowledge and Social Imagery*. London: Routledge and Kegan Paul.
- Callon, Michel, and Bruno Latour. 1992. "Don't Throw the Baby Out with the Bath School!" Pp. 343-68 in *Science as Practice and Culture*, edited by A. Pickering. Chicago: University of Chicago Press.
- Collins, H. M. 1975. "The Seven Sexes: A Study in the Sociology of a Phenomenon, or the Replication of Experiments in Physics." *Sociology* 9 (2): 205-24.

- . 1981a. "Son of Seven Sexes: The Social Destruction of a Physical Phenomenon." *Social Studies of Science* 11:33–62.
- . 1981b. "Stages in the Empirical Programme of Relativism." *Social Studies of Science* 11:3–10.
- . 1981c. "What Is TRASP: The Radical Programme as a Methodological Imperative." *Philosophy of the Social Sciences* 11:215–24.
- . 1983a. "The Sociology of Scientific Knowledge: Studies of Contemporary Science." *Annual Review of Sociology* 9:265–85.
- . 1983b. "An Empirical Relativist Programme in the Sociology of Scientific Knowledge." Pp. 85–114 in *Science Observed: Perspectives on the Social Study of Science*, edited by K. D. Knorr-Cetina and M. Mulkay. London: Sage
- . 1984. "Concepts and Practice of Participatory Fieldwork." Pp. 54–69 in *Social Researching*, edited by C. Bell and H. Roberts. London: Routledge & Kegan Paul.
- . 1988. "Public Experiments and Displays of Virtuosity: The Core-Set Revisited." *Social Studies of Science* 18:725–48.
- . (1985) 1992 *Changing Order: Replication and Induction in Scientific Practice*, 2d ed. Chicago: University of Chicago Press.
- . 1994. "A Strong Test of the Experimenters' Regress." *Studies in History and Philosophy of Science* 25 (3): 493–503.
- Collins, H. M., and Trevor J. Pinch. 1982. *Frames of Meaning: The Social Construction of Extraordinary Science*. London: Routledge & Kegan Paul
- Collins, H. M., and Steven Yearley. 1992. "Epistemological Chicken." Pp. 301–26 in *Science as Practice and Culture*, edited by A. Pickering. Chicago: University of Chicago Press.
- Dickson, C. A., and Bernard F. Schutz. 1995. "Reassessment of the Reported Correlations between Gravitational Waves and Neutrinos Associated with SN 1987A." *Physical Review D* 51:2644–68
- Ferrari, V., G. Pizzella, M. Lee, and J. Weber. 1982. "Search for Correlations between the University of Maryland and the University of Rome Gravitational Radiation Antennas." *Physical Review D*, ser. 3, 25 (10): 2471–86.
- Festinger, Leon, H. W. Riecken, and S. Schachter. 1956. *When Prophecy Fails*. New York: Harper.
- Fleck, Ludwik (1935) 1979. *Genesis and Development of a Scientific Fact*. Chicago: University of Chicago Press.
- Franklin, Alan. 1994. "How to Avoid the Experimenter's Regress." *Studies in the History and Philosophy of Science* 25 (3): 463–91
- Gallson, Peter. 1987. *How Experiments End*. Chicago: University of Chicago Press.
- Gieryn, Thomas. 1983. "Boundary-Work and the Demarcation of Science from Non-Science: Strains and Interests in Professional Ideologies of Scientists." *American Sociological Review* 48:781–95.
- Gooding, David. 1985. "In Nature's School: Faraday as an Experimentalist." Pp. 104–35 in *Faraday Rediscovered: Essays on the Life and Work of Michael Faraday, 1791–1876*, edited by D. Gooding and F. A. James. London: Macmillan.
- Holton, Gerald. 1978. *The Scientific Imagination*. Cambridge: Cambridge University Press.
- Infeld, Leopold. 1941. *Quest: The Evolution of a Physicist*. London: Gollancz.
- Knorr-Cetina, Karin. 1981. *The Manufacture of Knowledge*. Oxford: Pergamon.
- . 1991. "Epistemic Cultures. Forms of Reason in Science." *History of Political Economy* 23:105–22.
- Kuhn, Thomas S. 1961 "The Function of Measurement in Modern Physical Science." *ISIS* 52:162–76.
- . 1962. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

- Latour, Bruno, and Steve Woolgar. 1979. *Laboratory Life: The Social Construction of Scientific Facts*. London: Sage.
- Lynch, Michael. 1985. *Art and Artifact in Laboratory Science: A Study of Shop Work and Shop Talk in a Research Laboratory*. London: Routledge & Kegan Paul.
- MacKenzie, Donald. 1981. *Statistics in Britain, 1865-1930*. Edinburgh: Edinburgh University Press.
- Merton, Robert K. 1942. "Science and Technology in a Democratic Order." *Journal of Legal and Political Sociology* 1:115-26.
- . 1963. "Resistance to the Systematic Study of Multiple Discoveries in Science." *European Journal of Sociology* 4:250-82.
- . 1973. *The Sociology of Science: Theoretical and Empirical Investigations*. Chicago: Chicago University Press.
- Mulkay, M., J. Potter, and S. Yearley. 1983. "Why an Analysis of Scientific Discourse Is Needed." Pp. 171-203 in *Science Observed: Perspectives on the Social Study of Science*, edited by K. D. Knorr-Cetina and M. Mulkay. London: Sage.
- Pickering, Andrew. 1981. "Constraints on Controversy: The Case of the Magnetic Monopole." *Social Studies of Science* 11:63-93.
- . 1984. *Constructing Quarks: A Sociological History of Particle Physics*. Edinburgh: Edinburgh University Press.
- Pinch, Trevor J. 1981. "The Sun-Set: The Presentation of Certainty in Scientific Life." *Social Studies of Science* 11:131-58.
- . 1986. *Confronting Nature: The Sociology of Solar-Neutrino Detection*. Dordrecht: Reidel.
- Preparata, Giuliano. 1990. "'Superradiance' Effect in a Gravitational Antenna." *Modern Physics Letters A* 5 (1): 1-5.
- Rosenthal, Robert. 1978. "Interpersonal Expectancy Effects: The First 345 Studies." *Behavioural and Brain Sciences* 3:377-415.
- Shapin, Steven. 1979. "The Politics of Observation: Cerebral Anatomy and Social Interests in the Edinburgh Phrenology Disputes." Pp. 139-78 in *On the Margins of Science: The Social Construction of Rejected Knowledge*. Vol. 27 of *Sociological Review Monograph*, edited by R. Wallis. Keele: Keele University Press.
- . 1994. *A Social History of Truth: Civility and Science in Seventeenth-Century England*. Chicago: University of Chicago Press.
- . 1995. "Here and Everywhere: Sociology of Scientific Knowledge." *Annual Review of Sociology* 21:289-321.
- Shapin, Steven, and Simon Schaffer. 1987. *Leviathan and the Air Pump: Hobbes, Boyle, and the Experimental Life*. Princeton, N.J.: Princeton University Press.
- Travis, George D. L. 1981. "Replicating Replication? Aspects of the Social Construction of Learning in Planarian Worms." *Social Studies of Science* 11:11-32.
- Traweek, Sharon. 1988. *Beamtimes and Lifetimes: The World of High-Energy Physicists*. Cambridge, Mass.: Harvard University Press.
- Weber, J. 1984. "Gravitons, Neutrinos and Antineutrinos." *Foundations of Physics* 14 (12): 1185-1209.
- Weber, J., and B. Radak. 1996. "Search for Correlations of Gamma-Ray Bursts with Gravitational-Radiation Antenna Pulses." *Nuovo Cimento B* 111 (6): 687-92.
- Winch, P. G. 1958. *The Idea of a Social Science*. London: Routledge and Kegan Paul.
- Wittgenstein, L. 1953. *Philosophical Investigations*. Oxford: Blackwell.

Making Use of the Past: Time Periods as Cases to Compare and as Sequences of Problem Solving¹

Jeffrey Haydu

University of California, San Diego

This article examines methodological issues that arise when using information from one historical period to illuminate another. It begins by showing how the strengths and weaknesses of methods commonly used to compare institutions or regions reappear in comparisons between times. The discussion then turns to alternative approaches. The use of narrative and of path dependency to construct explanatory sequences are strategies that strike a welcome balance between causal generalization and historical detail. But these approaches typically fail to identify either the causal mechanisms or the trajectories that link events in different eras. These gaps can be filled by rethinking sequences of events across periods as reiterated problem solving. Successive U.S. industrial relations regimes since 1900 are used to illustrate this methodological strategy.

Wise parents, lecturers in history, and reflective journalists join historical sociologists in urging us to study our past, the better to understand our present. Parents captivate their children with tales of "when I was your age," drawing cautionary lessons from the old days to guide contemporary youth. Professors continue to invoke Santayana on the perils of ignoring the past, implying that careful study of previous epochs can usefully guide policy today. And journalists draw parallels between current developments—rapacious monopolists, widening income gaps, rogue world leaders, office automation—and their ostensible predecessors, as if previous outcomes forecast future developments.

In more or less self-conscious and systematic ways, these arguments simplify characteristics of different time periods, picking out and high-

¹ This article has benefited from the comments of Kathy Mooney, Colin Haydu, Howard Kimeldorf, Tien-Lung Liu, Mark Jones, Ákos Róna-Tas, Gershon Shafir, Clifford Staples, graduate students in my seminar on Comparative and Historical Sociology, and the *American Journal of Sociology*'s reviewers. Direct correspondence to Jeffrey Haydu, Department of Sociology 0533, University of California, 9500 Gilman Drive, La Jolla, California 92093-0533.

lighting some features of one period (e.g., the retreat from Progressive social engineering to *laissez-faire* in the 1920s) to juxtapose with selected features of another (the 1980s backlash against the welfare state). In this article, I consider *how* our understanding of one period may profit from juxtaposing it with another. Although consideration of the past to shed light on the present is the most familiar example, the question can be posed for any two periods: How do we use information about one to illuminate the other? I have two main goals. The first goal is to explore certain parallels between comparing times and comparing places; the second is to show how the use of narrative and path dependency models to link events between periods can and should be supplemented by attention to sequences of problem solving.

The first goal presupposes the validity of periodization. Sociologists and historians routinely divide the past into temporal chunks, although they often argue about the dates and characteristics that most usefully set one period off from another. These periods can be viewed as separate cases, and comparing them has much in common with comparing social institutions (such as welfare states or religions) or processes (such as revolutions or professionalization) that occur in different places. Although much has been written on the logics of comparing places and institutions, there has been no systematic exploration of how these logics apply to comparing periods. Exploring the parallels between comparing places and comparing times has at least two payoffs. One is simply to show the ways in which familiar strengths and weaknesses of variable-based or interpretive methods reappear in comparisons of different periods. The second payoff emerges as the analogy between comparing places and comparing times breaks down. When dealing with time periods, two cases can be related not only through comparisons and contrasts but *also* through sequences of events. This characteristic presents distinctive opportunities and difficulties and requires alternatives to conventional comparative methods.

The alternatives most commonly recommended rely on a self-conscious use of narrative, including the variety of narrative known as path dependency, to organize events into causal sequences. My second major goal in this article is to assess the separate merits of narrative and of path dependency for the specific tasks of establishing and explaining connections between events in different periods. Recent work urging the use of narrative in historical sociology points in the right direction (Somers 1996; Gotham and Staples 1996; Abbott 1992a; Isaac 1997). These studies advocate causal explanations that are sensitive to historical context and contingency, but they offer little guidance on how to develop such explanations. The narrative variant "path dependency" takes us several steps further. It specifies how contingencies steer historical change and what mechanisms "lock in" a given path. Thus, path dependency treats the idiosyncra-

sies of each historical period as resources for, rather than impediments to, causal explanation. But this approach also understates the influence of each historical turning point on later developments, and it obscures larger trajectories across periods.

We can remedy the deficiencies of conventional comparative methods by rethinking the connections between events in different time periods as reiterated problem solving. This heuristic device overcomes the limitations of interpretive comparisons by identifying and making use of continuities across periods, and it avoids certain pitfalls of variable-based comparisons by putting historical particulars to explanatory work. Organizing events into sequences of problem solving that span different periods also has advantages over narrative and path dependency. It provides a plausible way to represent and account for historical trajectories; it builds social actors and multiple causal time lines into the explanatory account; and it offers a richer sense of how earlier outcomes shape later ones. In the final section of the article, I use empirical examples from the history of U.S. industrial relations to illustrate this approach and to differentiate it from rival perspectives. I do not propose reiterated problem solving as a methodological Swiss army knife, handy for all historical chores. No such all-purpose tool exists. Instead, I offer it in the spirit of middle-range theory, as an approach well suited to some—but not all—types of social change.

COMPARING TIME PERIODS

The advantages of narrative, path dependency, and reiterated problem solving for comparing and connecting events in different time periods can best be appreciated against the backdrop of more conventional comparative methods. Most surveys of comparative historical sociology make an ideal-typical distinction between individualizing (sometimes labeled “interpretive”) and generalizing (or “variable-based”) comparisons (Abrams 1982; Skocpol 1984; Tilly 1984; Ragin 1987). The most widely accepted use of comparison falls in the latter category, seeking to identify causal relationships common across cases. The pursuit of causal generalizations, in turn, requires discounting the individuality of each case, abstracting selected variables from their historical setting, and making relationships among these variables the analytical focus.

In principle, advocates of the generalizing approach would join Przeworski and Teune (1970) in replacing scientifically irrelevant labels like “France” or “Germany” with dependent and independent variables that operate, like the Euro currency, without respect for nationality. In practice, comparativists in this genre differ in the level of generalization they deem appropriate for causal analysis and compatible with the integrity of each case (Skocpol 1984; Goldstone 1997; Goldthorpe 1997). Some limit

themselves to more or less narrowly defined and historically circumscribed phenomena like revolutions or fascism, and they caution against stretching causal findings too far beyond the cases studied (Skocpol 1979; Goldstone 1991; Brustein 1996). Within this narrowed range, however, the goal is still to abstract from particular phenomena (e.g., revolutions) characteristics and causal factors common across cases. Thus, Skocpol highlights the peasant rebellion and state collapse typical of social revolutions, including the Russian one, rather than examining how the peculiar contradictions of Russia's autocratic modernization led to revolution (cf. Skocpol [1979] with McDaniel [1988]). For other comparativists, even this level of generalization shows a deplorable lack of scientific ambition. Kiser and Hechter (1991) argue that sociologists ought to search for *general* causal laws. This requires that they conceptualize variables and causal relations among them at a level of abstraction suitable for universal applicability. State autonomy, for example, should be rethought in terms applicable to any setting in which there are rulers more or less dependent on their subjects for resources and subjects more or less able to monitor and sanction rulers—that is, virtually all known societies. Then, by making simplifying assumptions about rulers' motives, sociologists could develop testable propositions about the general causal relationships between state autonomy and state policy. If this approach results in "some loss of descriptive accuracy" (Kiser and Hechter 1991, 21 n. 45) about the interests of real actors in particular cases, that is merely the price to be paid for science (see also Goldthorpe 1997).

Individualizing comparisons treat idiosyncratic events and cultural nuances as essential *not* just for "descriptive accuracy" but for an adequate explanation of particular outcomes. By their very nature, these studies will yield conclusions of limited applicability.¹ In part, this reflects the practical constraints of labor-intensive research, which allows mastery of, at most, a handful of cases. But the underlying approach of individualizing comparisons also discourages generalization. Where the goal is to identify how social actors make sense of their worlds and to uncover the local cultures that guide them, researchers are likely to insist on the idiosyncratic character of their subject matter. Sociologists concerned with impersonal structural forces and causal determinants may be equally inclined to treat each case as a historical individual, arguing that no other society displays the same precise configuration of causal influences. Why, then, bother with comparison? In this style of comparative historical sociology, juxtaposing cases serves to highlight important differences and to discipline causal analysis. Lamont (1992), for example, pinpoints the con-

¹ The best example is the work of Reinhard Bendix, from *Work and Authority* ([1956] 1974) to *Kings or Peoples* (1978).

trasting languages of social exclusion by playing French and American professionals off against each other; Fredrickson (1988), by comparing American and South African race relations, highlights the causal importance of liberal ideology in the abolition of slavery and Jim Crow. Certain theoretical generalizations are at work here. Lamont frames her study with reference to Bourdieu's analysis of cultural capital, while Fredrickson uses ideal-types of racial exploitation. But these are starting points rather than goals for research. The authors do not intend the cultural meanings or causal configurations thus identified to be broadly applicable.

Strategies for using the past to illuminate the present often replicate these generalizing or individualizing styles of comparative sociology. With two periods as two cases, some would use comparison mainly to tease out important contrasts. Others would look for causal patterns that hold across temporal settings. As in the comparison of different institutions or places, these approaches to comparing different time periods have their characteristic virtues and vices. The balance sheet for generalizing approaches is the more complex of the two. I will begin with that.

Different Times, Same Causal Relationships

One way to compare temporal cases follows the logic of Mill's methods of difference and agreement. Ever since Galton's critique of Tylor in the late 19th century, methodologists have pointed out that Mill's experimental logic is inappropriate for proving historical explanations (Lieberman 1987; Burawoy 1989). Yet this logic continues to guide the selection and conceptualization of cases and to provide rules of thumb in the "detective work" (Goldstone 1997) undertaken by many scholars. Lipset (1977, 1986, 1991), for example, loosely periodizes working-class politics and unionization and claims to find the same causal pattern (in which America's liberal culture undermines popular support for statist politics and collectivism) in each. Skocpol applies the method of difference in periodizing the history of American public welfare. Her contrasts between the post-Civil War pension system and the New Deal welfare state highlight differences in critical variables between periods (such as the presence or absence of reform movements well adapted to America's political system). She then invokes these contrasts to support a more general causal account of welfare policy (Skocpol 1992).

Mill's deficiencies for validating comparative causal arguments are no less apparent when scholars seek causal regularities across time periods. In part, the problem is one of oversimplification. Any construction and comparison of cases, of course, abstracts some characteristics from complex settings and thus suppresses historical detail. The degree of simplification deemed legitimate is largely a function of intellectual interests and

tastes. But where the goal of cross-period comparison is causal generalization, there is an additional bias against historical particulars. The unique features of each temporal case may offer local color, but they can only be distractions from the explanatory business at hand.³ A common criticism of Skocpol's theory of revolution targets her assumption that a sound causal account must be applicable to all cases; this assumption leads her to discount sources of revolution, such as Russia's working class, that appear only in single cases (Burawoy 1989). The same assumption applied to cross-period comparisons similarly ignores period-specific influences and causal configurations. Wedded to his encompassing explanation for union weakness in U.S. history, Lipset (1986) attributes union decline in the current period to resurgent individualism. Comparing time periods in this way blinds him to the possibility that recent union decline reflects a conjuncture of long-standing and uniquely contemporary influences—that an interaction between America's distinctive industrial relations and global economic pressures is leading employers to repudiate the New Deal system.⁴

The search for consistent causal patterns across time periods imposes some additional costs, ones that may be less apparent in positivist comparisons of different social units. These are liabilities, I should add, only if one takes seriously the idea that there *are* time periods. Dividing history into meaningful sections involves both historiographical conventions and theoretical judgments about what constitutes a more or less unified "age," how that period differs from others, and where to locate the boundaries between periods. These conventions and judgments are mutable. Still, few students of history would accept the very different convention of treating the past as a series of data points in a seamless and uniform continuum of time. Instead, they organize time into chunks with defining themes, key events, and prevailing constellations of social forces. If it is granted that *some* such periods are useful constructs, then using the past to increase the *N* available for causal analysis involves two methodological dilemmas. First, period comparisons run afoul of the requirement that successive tests of a causal hypothesis must be both equivalent (approximating a controlled experiment) and independent (the results of one trial must not influence another). Second, period comparisons of this type cannot account for changing causal patterns across periods. Accounting for those changes demands the use of alternative methodological strategies.

³ Discounting historical detail, of course, makes the historical sociologist's research a good deal easier. Kiser and Hechter, e.g., commend general theories with "high analytic power" for their "low data input requirements" (Kiser and Hechter 1991, p. 10).

⁴ This is a historical example of the more general point, made by Liebersohn (1987) among others, that the scientific model emulated by many sociologists leads them to rule out different causes for similar outcomes.

As Sewell (1996) points out, in comparative analysis, the independence and equivalence of cases tend to be mutually exclusive. The more independent two cases are in space and time, the less likely they are to be equivalent. The same objection applies with greater force where the cases are two time periods in a single society. The charge that time series analysis is ahistorical (Isaac and Griffin 1989) highlights the lack of equivalence between periods. As social settings change, so do causal relationships. There is no good reason to assume that findings from one period support causal claims for another period. In quantitative work, the perils of this assumption are suggested by Collier and Mahoney (1996), who note how a reliance on any particular time span can lead to selection bias and conclusions that fail to fit a different period. Criminologists who have promoted new or revised causal accounts to fit the recent decline in crime rates, for example, may find their explanatory successes short-lived if the drop in crime proves anomalous.⁵ The same lack of equivalence can undermine causal arguments of a more qualitative sort. The relationship between work experiences and class politics claimed by Marx and many of his followers had some clear explanatory value for late 19th-century European and North American industrial societies. But this relationship was fundamentally reconstructed by changes in residence and consumption patterns in the early 20th century and by changes in political economy ushered in by World War I (Haimson and Tilly 1989). The assumption of independence is no less faulty. It would be hard to argue persuasively that later periods are unaffected by earlier ones, even if we disagree about what those effects are. For most of the 1970s and 1980s, for example, cigarette smoking declined among teenagers, suggesting the salutary effects of public health campaigns. Since 1992, teenage smoking has edged back up. A plausible hypothesis, consistent with the assumption of causal regularity, is that cutbacks in antismoking advertising are the culprit.⁶ Another possibility, however, is that the very success of ad campaigns and school programs in presenting nonsmoking as the sensible choice has made cigarettes a more attractive vehicle for teenage rebellion. An outcome of the cigarette wars in one period, then, may have not only eliminated the prior effects of antismoking campaigns but also set in motion a very different causal dynamic. A methodology that requires independence

⁵ For a comparable gaffe in assessing the relationship between speeding and traffic fatalities, see Campbell and Ross (1968).

⁶ This is the flip side of Lieberman's asymmetrical causality. Much as assumptions of symmetry may lead social scientists to rule out a causal factor if its *second* change in value has no impact on the dependent variable, so a reversal in the dependent variable may lead social scientists to mistakenly invoke whatever factor appeared responsible for the initial changes.

between cases can thus be applied to different time periods only by disregarding the influence of the past on the present or by choosing periods so remote in time as to make the comparison dubious.

Causal models spanning two or more periods do not need to deny that these periods differ. Existing differences, however, reflect varying values of the causal factors in the explanatory model, not alterations in the model itself. If strike rates vary with labor market conditions, then periods of labor surplus will have fewer strikes, and periods when jobs go begging will have more. But what of times when this pattern itself does not hold? Here arises a second methodological dilemma for temporal comparisons devoted to causal regularities: given their own standards for what constitutes a causal explanation, they cannot explain a change in pattern between periods. The *effects* of these shifts certainly can be represented in statistical models. Dummy variables can be used to show that period-specific influences do make a difference. This is Skeels's approach. His study of the determinants of strikes (Skeels 1982) includes dichotomous dummy variables for the Wagner Act, Democratic Party power, the Great Depression, and the period 1921–29 (to “capture the effects of the open-shop drive, welfare capitalism, and some other events of the period” [p. 499]). The independent contribution of these dummies to variation in strike rates can then be measured. For those interested in the history of American labor, however, the use of dummy variables is at best a halfway house for recovering positivists. When shifts in employer policies, economic conditions, electoral politics, and labor law are regarded as important parts of the story of *how* strike patterns have changed, it is necessary to explain those shifts, not merely register their impact. A method that relies on abstracting variables from their temporal context in order to discover causal regularities, however, cannot explain events or circumstances specific to one time. Worse, it cannot even represent, much less explain, the links between these events. Skeels's work, for example, cannot engage the ways that employer policies contributed to economic depression, how these two together fueled political realignment, and how all of these factors combined to stimulate labor law reform. Accounting for this sort of historical *trajectory* demands very different strategies for connecting “data” from different periods.

Different Times, Different Stories

Like analyses of causal regularities, individualizing comparisons have their own distinctive strengths and weaknesses. These appear as clearly when the cases are times as when they are places. Comparison in this genre highlights the distinctive features of each period rather than variables and causes common across time. Individualizing comparisons thus

yield accounts that are descriptively richer and that treat historical details as resources for interpretation rather than as obstacles to causal analysis. But this also sharply limits the *explanatory* value of historical studies for the present. Studies of this nature, Sean Wilentz maintains in a recent editorial, teach us "the pastness of the past—and with it the humanizing lesson that people in history, although every bit as intelligent as ourselves, thought and lived in ways very different from our own" (Wilentz 1997). Even Bendix draws broader lessons from the past. Best known for his cautious cross-national comparisons, he applies the same logic to time periods. In *Work and Authority in Industry*, for example, Bendix contrasts managerial ideologies in the 1920s and 1930s with those characteristic of the 1910s. Rather than viewing workers as motivated by shortsighted goals, personnel managers in the interwar years emphasized the need to nurture employees' inner feelings and social needs (Bendix 1974, pp. 293–97). Bendix never denies that some things remained much the same over the two periods. For all their loving attention to worker psychology, employers denounced unionism as belligerently in the early 1930s as they did 20 years earlier. Bendix's research interest, however, lies in the changing justifications for managerial authority. Given that interest, comparison of time periods serves to tease out differences and highlight shifts that might otherwise have gone unnoticed.

Individualizing comparisons are neither atheoretical nor, necessarily, grounded in historical minutiae. Michael Burawoy sets the stage for *Manufacturing Consent* with an individualizing comparison of feudal and capitalist labor processes. In 10 pages, there is one passing reference to a place (England) and two to times ("prior to the Factory Acts" and "nineteenth-century capitalism"). The comparison nonetheless serves Burawoy's purpose of clarifying an essential dilemma of the capitalist labor process: How do employers simultaneously secure and obscure surplus value (Burawoy 1979, pp. 20–30)? The logic of Bendix's inquiry is similar, although the level of historical detail is not. He begins with Weber's conceptualization of authority relations and the problem of legitimacy, and these guide his historical research. Like Burawoy, the generalizations he does allow take the form of recurrent dilemmas (rulers need to justify their power to themselves and, often, to the ruled), not timeless variables or causal relationships. Finally, the assumption Bendix makes at the outset is that this dilemma will be resolved in different ways at different times, thus leading him back to the particularities of individual cases.

When we use one time period to help explain another, two questions arise: How do we account for the differences we identify? And how did we get from there to here? Individualizing comparisons are poorly designed to answer either. Bendix, of course, has plausible explanations for why human relations management appeared in the 1930s, much as Lamont (1992)

has plausible explanations for why American and French upper-middle-class males draw different kinds of status distinctions. The problem is that in constructing their comparisons of times and places to highlight differences, individualizing comparativists must choose cases that provide them with no leverage for causal analysis. Explanation is still possible, but not through a comparison of the selected cases. Instead, scholars typically recount the history of each single case to explain (or at least to describe) the origins of a given trait or outcome. In methods based on the analysis of causal regularities, by contrast, the characterization of cases and the explanation for similarities and differences are both part of the same logic of comparison, however deficient these methods may be in other respects.

In answering the second question—How are phenomena at time 1 connected to those at time 2?—individualizing comparisons are no more useful than generalizing ones. They are certainly better suited for *identifying* the differences between two periods. Close comparison of the constituencies and spatial distribution of popular protest in Paris in 1848 and 1871 suggests to Gould (1995) that fundamentally different identities served as the basis for mobilization. In 1848, artisans mobilized as members of a working class; in 1871, neighbors did so as citizens of Paris. The contrast sets up the task of showing how identities were transformed. But at this point, the initial comparison has served its purpose and can offer no leverage for establishing causal claims. It is not on the basis of comparing 1848 and 1871, for example, that Gould chooses to focus on changing social networks as the source of new identities; he has independent theoretical reasons for pursuing that lead. Nor does the comparison itself point to the key steps (urban renovation and the changing geography of industry) that link 1848 and 1871 into a coherent journey. To present and explain these connections, scholars like Gould and Bendix abandon their individualizing comparisons and weave stories.

NARRATIVES, PATHS, AND PROBLEM SOLVING

Much as with the comparison of places, then, interpretive and variable-based comparisons of time periods face tensions between an appreciation for the individuality of each period and attention to regularities across cases. Moreover, neither comparative logic can bring contrasts and continuities together in a coherent explanation of the connections between periods. But here the analogy between comparing times and comparing places breaks down. Both kinds of comparisons can make fruitful use of differences and similarities. Time periods, though, normally can *also* be connected through sequences of events—sequences that may result in one case being transformed into another. This feature of temporal cases creates distinctive opportunities and difficulties for which traditional com-

parative methods are of little use. How do we rethink "cases" when one case becomes another over time? How do we select and organize events into explanatory rather than merely temporal sequences?⁷ And how do we validate explanatory accounts once "explanation" is rethought as a matter of demonstrating *how* past events led to later outcomes rather than as a matter of applying covering laws to new cases? By focusing on events, arranging them in temporal order, and asking how sequences are also causal chains, we have an agenda for moving beyond methods that merely compare periods. Accomplishing these tasks would also help answer the call for sociological explanations that recognize historical contingency, multiple and mutable patterns of causality, and the causal importance of temporality itself (Sewell 1996; Somers 1996; Aminzade 1992). It is a compelling program. How can we realize it?

The strategy generally recommended for constructing properly historical explanations treats sequences as narratives. In the following section, I evaluate this alternative, along with one of its variants, path dependency. Narrative and path dependency are valuable scholarly tools. They help impose both temporal and explanatory order on events; they respect historical context and contingency within each period without foregoing causal explanation; and they accomplish both of these ends in part by invoking forms of causality that are better suited to historical change than are the covering laws of Kiser and Hechter or Skocpol (Quadagno and Knapp 1992; Somers 1996). Nevertheless, narrative and path-dependent explanations, as advocated and often as practiced, fall short when confronted with the task of explaining connections between events in different time periods. Advocates of narrative accounts are usually vague about the causal mechanisms that make sequences of events across periods *explanatory* sequences. One major virtue of path dependency is that it specifies some of these mechanisms. This particular form of narrative, however, also impoverishes our understanding of larger trajectories. As an alternative to both the vagaries of narrative and the narrowness of path dependency, I advocate connecting events between periods through sequences of problem solving. Models of reiterated problem solving are a kind of narrative. And like path dependency, reiterated problem solving constructs narratives of historical switch points that are followed by more or less durable social regimes. Both of these approaches, moreover, make

⁷ Sequences in which there are no such explanatory links can occur. T. H. Marshall's (1964) account of civil, political, and social rights arranges these in sequence, but it does not assume that the development of civil rights is one cause of later gains in political rights. Even in classic evolutionism, prior stages do not cause later ones. Instead, some overarching logic (such as "differentiation and integration") is held to explain the succession of stages.

legacies from earlier historical junctures an important part of causal explanation. Where reiterated problem solving differs from path dependency is in the way it links these historical junctures into coherent sequences and in its account of how earlier historical turning points shape later ones.

Connecting Periods through Stories

Narrative has been widely prescribed as a cure for much that ails historical sociology. It promises to rejuvenate the study of class (and other group) formation, calling attention to how social actors construct meaningful stories of individual and collective identities by weaving together interpreted events (Steinmetz 1992; Somers 1992). Narrative is also a useful part of a historical sociologist's tool kit, handy for the construction of temporal and explanatory sequences (Griffin 1992; Maines 1993; Gotham and Staples 1996). The analyst organizes events into a story with a beginning, middle, and end, central characters, and a coherent plot. Once there is a story line, events at one point in time can be "explained" with reference to prior plot developments. The analyst-as-storyteller identifies the "inherent logic" that makes one event follow from another (Abbott 1992a, p. 445; Griffin 1993; Isaac 1997). This storytelling approach to explanation, proponents assert, is the best way to represent how causal relations are embedded in particular contexts and enacted over time. Narrative also comfortably accommodates the inescapable contingency of historical sequences, whereby events alter the direction of social change and transform social structures (Sewell 1996; Abbott 1997).

For the purposes of using elements from one period to explain another, calls for narrative are more effective as critiques than as guides. What, besides the seductive charms of a good yarn, persuades us that narrative links are causal determinants? One strategy for testing those links within a single time period is event-structure analysis (Griffin 1993). Event-structure analysis forces observers to be more methodical in identifying which events supposedly "led to" later ones.⁸ Further, event-structure analysis helps build the case that these connections are causal rather than merely sequential connections by using counterfactuals to "test" assertions that Y occurred *because of* X. This strategy thus promises to combine causal analysis with attention to unique sequences of events.

But what of connections between events *across* periods? Abbott suggests that narrative can do this job, too. A narrative approach can identify

⁸ Stinchcombe's (1978) advice to decompose historical sequences into causal bits discovered through deep analogies represents a similar strategy.

common patterns in sequences of events in two different periods (as in the "natural history" of revolution) or it can treat elements of each period as themselves events in a larger and longer sequence (as in stages of professionalization) (Abbott 1983, 1992a; Abbott and Hrycak 1990). The first of these strategies surely adds rigor to impressionistic judgments that certain patterns of events (e.g., careers) repeat themselves. But like the quest for causal regularities, picking out common narrative structures fails to establish connections *between* eras. Formal parallels in sequences of events within two separate periods can tell us nothing about how events in one influence those in the other.⁹ If, instead, the contents of two periods are themselves treated as events in a larger narrative sequence, then characteristics of the present may be explained with reference to elements of the past: both are understood as chapters in an ongoing story. Before we accept the claim that one chapter really is connected to and accounts for a later chapter, however, additional work needs to be done. Much as with statistical correlations that aspire to be causal relations, narrators who want to interpret sequential events as parts of a coherent story must identify the mechanisms through which characteristics or events at time 1 lead to or are transformed into characteristics or events at time 2. This is relatively easy when the story begins and ends within a single epoch. In this setting, there are historical actors to help convey the story and link the consequences of events at one time to the conditions shaping events at another time (Hall 1995; Rueschemeyer and Stephens 1997). When the narrative spans time periods, however, individual chapters are located in different temporal contexts, making it at once more important and more difficult to pinpoint the mechanisms that transmit the influence of one sequence to the next. Identifying how these transmissions occur is important in order to avoid teleological arguments in which some overarching logic of development, divined by the observer, is all that ties together historical events or supports causal links (cf. Sewell [1996] on "teleological temporality"). The task is difficult because, as the gap between eras widens, causal carriers become less obvious and more difficult to isolate. In adjacent time periods, for example, cohorts may convey the influence of events. Older skilled workers brought memories of shop-floor militancy from the World War I era into the ranks of semiskilled factory workers in the 1930s; similarly, participants in San Francisco's 1934 general strike kept the syndicalist torch burning into later decades (Brody 1993; Kimeldorf 1988). Even as the interval between periods grows, transmission still

⁹ Abbott notes one methodological advantage in comparing narrative structures *versus* comparing relations among variables in different times. The fact that the two periods do not meet the experimental standard of independence compromises causal analysis but not narrative interpretation (Abbott 1992b, p. 73).

occurs through human beings. But the analytical focus has to shift from living individuals to social mechanisms that reproduce, for example, "traditions."

Connecting Periods through Paths

Path dependency offers a more rigorous way to identify these social mechanisms. Like other forms of narrative, path dependency organizes events and circumstances into temporal sequences. It makes these into *explanatory* sequences by identifying choices or conditions that foreclosed options and steered history in one or another direction (David 1986; Arthur 1988; North 1990). Path dependency narratives begin with a historical fork in the road; identify the turn taken and emphasize how subsequent developments make that choice irreversible. Piore and Sabel, for example, point to alternative methods for organizing industrial manufacturing in the early 20th century. Once American business adopted mass production rather than flexible specialization, industrial technology followed a developmental course from which departures became increasingly difficult (Piore and Sabel 1984). Voss identifies a turning point in American working-class politics in the late 1880s. At that time, the defeat of the Knights of Labor eliminated the institutional infrastructure and ideological visions that had supported class solidarity. The labor movement subsequently proceeded along tracks laid down by the sectional organization and business unionism of the American Federation of Labor (Voss 1993).

What sets path dependency apart is not so much the way it describes historical change as the way it explains those changes.¹⁰ When accounting for historical turns down one road rather than another, path dependency emphasizes contingency. Roy asserts that in the 1830s, the balance between public and private control of the corporation was still unsettled. Historical accidents, such as the timing of canal construction and railroad development, steered policy toward privatization. And because these contingencies varied from one state to another, so did the details of policy outcomes (Roy 1997). As against causal generalization, such path-dependent accounts thus incorporate period-specific, nongeneralizable causes. In explaining subsequent continuities, by contrast, path dependency offers a much more deterministic account. Compared to most calls for narrative

¹⁰ The term "path dependency" is often applied loosely to any claim that historical options or causal relations are constrained by prior developments. By this standard, reiterated problem solving is a subset of path dependency. But it is the more specific arguments about turning points and lock-in mechanisms that really make path dependency something more than plain historical common sense. My focus is on these arguments.

explanation, path dependency clearly identifies mechanisms that reproduce the outcome of a historical turning point. Conventional narratives link events at different times by placing them both in a common setting between the beginning and end of a coherent plot. Path dependency, on the other hand, depicts events as stations along a historical track, and it goes on to specify mechanisms that *keep* history on track. In David's (1986) classic account of the QWERTY typewriter keyboard, these factors include sunk costs in human and infrastructural capital. In Roy's (1997) study of the corporation, it is the power of a new corporate elite that ensures that contingent outcomes become durable ones.¹¹ If at turning points history is full of serendipity, along ensuing paths it becomes a steamroller.

In constructing explanatory sequences, path dependency thus replaces story lines with eventful and underdetermined switch points, on one side, and deterministic "lock-in mechanisms" on the other (Arthur 1988; Róna-Tas, 1998). This alternative to the vagaries of narrative has its costs, however. Advocates of narrative fail to clearly identify the mechanisms that link events into overarching tales. Path dependency identifies some of these mechanisms but fails to provide the overarching tales. In part, this is because discussions of path dependency rarely deal with *multiple* switch points that form more encompassing sequences. More importantly, the very nature of path-dependent explanations obscures larger trajectories across periods. Choices at each critical juncture nudge history down paths that then become difficult to escape. But those choices are themselves more or less accidental and, in North's (1990) account, exogenous. They bear no systematic relationship to subsequent turning points. This explanatory model reflects the original purpose of path dependency: to account for the stubborn persistence of suboptimal technologies or economic arrangements despite neoclassical expectations. But it leaves unappreciated the many ways that history's switchmen come along for the ride. Choices in one period not only limit future options, they may also precipitate later crises, structure available options, and shape the choices made at those junctures.

Connecting Periods through Problem Solving

My discussion thus far has generated a wish list for what should be accomplished when using events in one period to explain outcomes in another.

¹¹ Although Abbott continues to use the term "narrative" in his recent discussion of turning points and trajectories, he, too, emphasizes that this way of analyzing sequences is quite different from their interpretation as stories. Compare Abbott (1997) with Maines (1993) or Somers (1992).

As in the comparison of places, a balance should be struck between identifying continuities across periods and recognizing the distinctive character of each. Analysis of events or data from two or more periods should also help the observer *explain* differences between them. This is a task for which both the positivist search for causal regularities and individualizing comparisons between epochs are ill-suited. To be satisfactory, explanations should respect historical time by casting causal analysis in the form of sequenced events, with earlier happenings leading to and accounting for later ones. Explanations of this kind should, moreover, carefully specify the mechanisms through which causal influence is conveyed through time. Advocates of narrative fail to do so. Instead, an abstracted story line is all that connects successive events in an imputed causal chain. Finally, causal processes should do more than enable past events to switch historical directions. They should carry the explanatory weight of the past forward to shape later turning points and critical choices. Analyzing events from different periods in this way would offer a fuller understanding of historical trajectories than models of path dependency can provide.

One way of meeting these needs is to link facts from different periods into larger sequences of problem solving. Periods are demarcated on the basis of contrasting solutions for recurring problems, not different values of a causal variable or diverging outcomes between historical turning points. Continuities across temporal cases can be traced in part to enduring problems, while more or less contingent solutions to those problems are seen as reflecting and regenerating the historical individuality of each period. One of the main explanatory goals is to account for why at a given time human beings pursued one solution rather than another. A solution sets a new historical direction, and that developmental track limits future choices. This basic picture resembles path dependency. Beneath this picture, however, lies a very different explanatory model. A problem-solving approach, unlike a path-dependent account, would attend to the ways in which outcomes at a given switch point are themselves products of the past rather than historical accidents. Solutions may embody contradictions that generate later crises, and they bequeath tools and understandings with which later actors confront those crises.

What sorts of subjects might fall into the category of enduring problems? They may be dilemmas deduced from theory, as Bendix (1970, 1974, 1978) advocates in his methodological essays and in his comparative work on private and public authority. Or, they may be issues embedded in empirical events. Fredrickson (1988), for example, begins with substantive historical interests (such as the aftermath of slavery) and puts at center stage specific social actors (plantation owners) grappling with particular social issues (including the problem of labor mobility). Either way, the

starting point for research is not the formulation of hypotheses. For example, rather than posing a theoretical question about how state autonomy influences government policy in different epochs and then reasoning out testable propositions, the researcher might ask how political elites have coped with the power and defined the rights of economic corporations. After documenting and ordering the various strategies adopted from one period to another, the researcher would explain variations across periods. This task would involve investigating (among other things) contrasts in how social actors constructed the problem and what solutions appeared to be realistic within each historical context (Roy 1997).

How can the "commonality" of problems across periods be defined and delimited? In contrast to the dictates of path dependency and of variable-based causal models, one criterion must be the social actors' own understandings. These individuals may not have defined the problem they confronted in the same terms as the historical sociologist does, but there must be *some* correspondence between the observer's conception of a recurring problem and the social actors' experiences of confronting common obstacles and devising ways to surmount them (whether or not those innovations succeed as planned). This requirement has at least two benefits. First, it imposes a salutary check on sociologists' theoretical imperialism. This is the tendency, illustrated by functionalists and Marxists alike, to define essential social problems with little reference to historical actors' points of view. Second, this requirement compels the researcher to build into his or her methodological strategy the agency of social actors as they define problems, devise solutions, and take action.¹²

What sorts of actors does this kind of approach presume? They need not be rational; that is, their definitions need not be accurate, their solutions calculated, or their actions instrumental. Nor are these solitary or homogeneous actors. It may be easier for the researcher (and the reader) to organize explanatory sequences around particular actors' responses to problems as those actors understand them, thus maintaining a consistent narrative voice. But a problem-solving account should include other actors, with their own definitions of problems and proposed solutions. Practices at particular times will reflect the clash of rival solutions, and one group's initial remedies will be modified in the face of another's backlash.¹³ These dynamics will complicate interpretation but not alter the

¹² This is also Abbott's recommendation (Abbott 1992a, p. 429; Abbott and Hrycak 1990, p. 147).

¹³ Thus one may "tell the story" of Islamic intellectuals' response to the West in terms of how they balanced a desire for modernization with an insistence on a *non*-Western developmental path, and how and why their strategies for achieving this balance varied from one time to another. The story would obviously be incomplete without close

logic of reiterated problem solving. In any case, the methodological injunction is not to prejudge these issues. Instead, it is to construct explanatory sequences on the basis of how contestants work with problems, tools, and options inherited from the past. Historical trajectories linking events across periods are seen as a matter of successive solutions. Each is to be explained in part by the nature of the crisis, by the way the problems are understood, and by the range of solutions that are open in both subjective and objective senses.

Reiterated problem solving, then, involves a different way of defining cases and constructing explanatory sequences. It also involves a different relationship *between* casing and explanation than exists in generalizing or individualizing approaches. Unlike the latter, reiterated problem solving deploys contrasting cases to develop explanations both for period-specific solutions and for differences between periods. As against variable-based temporal comparisons, historical particularities become resources for identifying the contours and roots of particular solutions. In reiterated problem solving, explanation does not *require* the suppression of each case's individuality. Nor is there any assumption that causal explanations will be found that apply beyond each period—whether in the conventional sense of explaining variance or in Lieberman's sense of deep underlying causes for "the thing itself" (Lieberman 1987). Instead, it is assumed that a *full* explanation for solutions adopted at one time will not be wholly applicable to another.

These characteristics of reiterated problem solving dissolve the tension between continuities and contrasts found in both generalizing and individualizing comparisons of time periods. Studies in the first of these traditions tend to use one explanatory device (covering laws) to account for common patterns across periods and another one to accommodate irrepressible differences between temporal cases (including ad hoc explanations to deal with "historical residuals") (Goldthorpe 1997). Analyzing continuities across and contrasts between periods as part of a coherent sequence of problem solving makes these two explanatory goals part of the same intellectual enterprise. Explanations based on individualizing comparisons are also unable to reconcile contrasts and continuities. Comparisons that are useful for identifying differences cannot also be pressed into explaining the origins of distinctive outcomes. Rethinking those outcomes as divergent solutions for common problems enhances the explanatory value of the comparisons. Period comparisons can now be used to zero in on plausible, period-specific influences on the ways problems are defined and attacked (Does the growing visibility of the homeless make

attention to divisions among the intelligentsia and conflicts with other social groups within and outside their countries.

us think about the poor in different ways?). Comparison of selected temporal cases (such as times, like the 1870s, of rising vagrancy) can also help check our causal hunches.

The weaknesses inherent in generalizing and individualizing approaches make them easy marks. Path dependency is a considerably more robust target and one with which reiterated problem solving shares some common ground. Both path dependency and problem solving offer similar narratives of history as proceeding from critical junctures to durable regimes, and they both link events from different periods into historical sequences with explanatory value. But representing continuities as recurrent problems rather than as persistent outcomes has other advantages. One is to provide a richer sense of fate. Solutions at one time may do more than foreclose future options—their only role in path dependency. They may also lead to and shape the switch points confronted by later generations, drawing the fault lines along which later crises will erupt and creating options for new solutions.¹⁴ Second, by organizing casing and explanation around problem solving, this approach puts social actors at center stage (as recommended by, e.g., Tilly 1997, p. 47; Sewell 1996). This focus is recommended neither for sentimental humanistic reasons nor for its conformity with theoretical fashion. It is sound methodology. Making actors the strategic pivots of historical sequences provides another way in which reiterated problem solving, as compared to path dependency, provides a fuller sense of the past's causal influence.¹⁵ The ways that people cope in one period affect both how their descendants diagnose the next crisis and what remedies they have available. Events in one period thus play a much larger role in structuring later turning points and choices than path-dependent models envision.¹⁶ Third, by making the links between events and temporal cases turn on human problem solvers, reiterated problem solving supplements path dependency, providing a way to integrate multiple registers of historical time (Abbott 1983; Sewell 1996). At any given switch point, it is through historical actors that the break-

¹⁴ Here my approach differs from Abbott's. He prefers to methodologically separate each historical turning point from prior trajectories, although he adds that turning points *may* be linked by some longer-term, overarching historical processes (Abbott 1997, pp. 93, 101).

¹⁵ Douglass North (1990) is rare among economists in acknowledging that inherited cultural lenses and ideological commitments shape rational choices. But he then limits this influence; culture and ideology account only for marginal adjustments in the institutions that define a path, and they matter only to the extent that markets are imperfect.

¹⁶ This emphasis on *sequences* of problem solving also differs from Bendix's approach. In his work, generic problems of authority are solved in different ways in different settings, but those solutions do not, in turn, foment or shape later problems.

down of old regimes, the changing social definitions of the problem, and the development of new tools for coping—each with its own historical rhythm—come together.¹⁷

Reiterated problem solving does not meet all our methodological needs. I advocate it, instead, as a valuable supplement. The usual response to intractable problems in grand theory, after all, is to embrace middle-range theories suited to the research problem at hand. There is a similarly sensible answer to the general methodological dilemmas raised when connecting events and circumstances across periods. Rather than trying to reason out the one true way to resolve these issues across all empirical settings, we can develop middle-range heuristic tools that serve our particular scholarly goals. The strategy recommended here is appropriate where there are *either* theoretical or empirical reasons to treat institutional practices as temporary (and possibly contentious) accommodations for recurrent dilemmas. In these settings, reiterated problem solving is a more useful form of narrative than path dependency for constructing explanatory sequences. If this hardly covers all topics of interest to historical sociologists, it is still widely applicable. McDaniel's (1996) recent account of the pathologies of Russian development, for example, interprets successive epochs of Russian history as reflecting a recurrent tension between the demands of modernization and Russia's enduring cultural infrastructure. In each epoch—including Tsarist, Communist, and post-Communist—"the Russian idea" blends with period-specific developments to yield a distinctive configuration of ideas, policies, and social contradictions.¹⁸ These, in turn, structure the practical problems of governance and modernization that have confronted rulers from Nicholas II to Boris Yeltsin, have shaped the assumptions they bring to these problems, and have defined their opportunities for reform.

A model of problem solving does not, however, require any such assumptions about functional needs or insoluble dilemmas. In surveys of the history of welfare states, for example, scholars commonly organize their accounts around the problem of the dependent poor and construct temporal cases based on characteristic government strategies. They then move on to causal arguments for how state resources, class alignments, and prevailing conceptions of the poor (each shaped in part by differently

¹⁷ Abbott also notes that social action links different trajectories (or "multiple emplotments"; Abbott 1977, p. 99). However, he uses this point to show how networks of trajectories reproduce social structure rather than as a device for constructing explanatory sequences.

¹⁸ A similar logic has artisanal workers in England, France, and the United States adapting traditions (such as corporatism or Republicanism) inherited from preindustrial epochs to meet new challenges under new conditions. The leading examples are Thompson (1963), Sewell (1980), and Wilentz (1984).

paced historical trajectories) favor specific policies at specific times.¹⁹ As the next section demonstrates, a similar logic can be applied to employer policies regarding “the labor problem.”

A CASE IN POINT: FIGHTING UNIONS IN AMERICA

Over the last 15 years, top managers have mounted a campaign of “restructuring” to shift work to temporary or part-time employees, subcontractors (preferably nonunion), and foreign operations. These measures do more than cut wages and benefits; they also give employers the flexibility to quickly slough off or reassign manpower as product markets or production methods change. These are freedoms union contracts often curb (Appelbaum and Batt 1994). As an article in *Business Week* concludes, “America’s most successful companies seem to have decided that a workplace compact is necessary only for their most valued workers” (July 10, 1995, p. 22). Even for this core of valued employees, however, the workplace compact need not involve union representation and contracts. Managers who have sought to overcome rigid work rules and adversarial bargaining by broadening job skills and involving employees in “team” approaches to improving productivity and resolving conflicts have often done so without union partnership (Kochan, Katz, and McKersie 1986; Craypo and Nissen 1993).

The most common explanation for this offensive against unions highlights the challenges and opportunities presented by globalization. International product markets are more competitive and volatile, allegedly forcing businesses to cut labor costs and restructure the labor process. Globalized production and capital mobility, in turn, give employers new weapons to wield against organized labor. The steep decline in union density since the early 1970s shows which side is winning. (The literature is summarized in Applebaum and Batt [1994] and Moody [1997]). Analysts with more historical awareness remind us that, in America, antiunionism is nothing new. Business acceptance of collective bargaining and the right to join “outside” labor organizations was unenthusiastic and spotty even during the New Deal. Both before and after that time, union busting has been a popular employer sport. In the hands of a social scientist like Lipset (1986), this argument explicitly takes the form of a variable-based temporal comparison; similar causes (e.g., individualism) in different periods have similar effects (antiunionism). And like temporal comparisons in general, the argument certainly makes use of the past—but not enough use.

¹⁹ Students of American welfare policy have often applied this logic in their research. See, e.g., Skocpol (1992), Orloff (1988), Fraser and Gerstle (1989), Gans (1995), and Katz (1996).

Highlighting general themes across historical cases offers little leverage for understanding specific contemporary practices. Adequate explanations for the open shop drive in the 1900s or the American plan in the 1920s, for example, cannot be transferred wholesale to the present. The challenges faced by current employers, their diagnoses of the labor problem, and the opportunities open to them in our recent political and economic climate all differ from the past and require a historically specific account of their strategic choices.

Making productive use of the past thus involves not abstracting similarities but asking how solutions have differed and why. Some of the challenges faced by contemporary employers—notably globalization and slowed growth in manufacturing—were little known to their counterparts in the 1920s, when domestic markets permitted rapid expansion. Employers during the World War I era, for their part, perceived unions as a potential threat to industrial government, and they had at least some justification for those fears. Today, management has the luxury of defining the union menace in purely economic terms, and employer strategies for curbing union power carry few political costs. Such differences signal the need for caution in applying lessons from the past. As I argued above, the methodological goal should be to treat outcomes in different eras as parts of an intelligible sequence and recognize that past problems and solutions have cumulative influences on later ones. I also reviewed two narrative strategies for doing so: path dependency and reiterated problem solving. How does each make use of the past to help understand contemporary antiunionism?

Path dependency helps focus our attention on the ways that employers' reactions to contemporary challenges conform to paths laid down long ago. Piore and Sabel (1984), for example, interpret antiunionism as mass production's last gasp: so long as managers remain committed to this model of workplace organization, they will respond to market volatility by further economizing on worker skills and pay. Piore and Sabel's hopes for change also follow the logic of path dependency. Having traveled the Fordist road for over 60 years, U.S. business may finally be pushed by globalization on to the other developmental track, that of "flexible specialization." In this upbeat scenario, it is simply good business to enhance worker skills, participation, and security *beyond* anything that unions could achieve under the old adversarial regime (see also Hirschhorn 1997). These interpretations showcase the two key explanatory devices that underlie path dependency: switch points and lock-in mechanisms. The pessimistic scenario highlights the constraints of past outcomes, while the optimistic one sees the current crisis as a historic opportunity to turn from one production regime to another. In both respects, this narrative of path dependency enriches our understanding of contemporary antiunionism.

Narrating employer policy as reiterated problem solving retains this emphasis on constraining paths and turning points. But it also presents the past as a more complex and potent determinant of contemporary anti-unionism. It does so, first, by revealing several ways in which the current crisis is itself the product of prior employer strategies. For path dependency, the past matters because a historical trajectory set long ago still molds business responses to the exogenous stresses known as "globalization." Reiterated problem solving interprets those stresses as in some sense domestic phenomena, expressions of contradictions embedded in the old industrial relations regime. Second, the model of reiterated problem solving represents employer responses to current economic challenges as a complex legacy of *multiple* historical paths and switch points. In the rest of this section, I will illustrate these contributions by looking more closely at successive turning points in U.S. labor relations.

A first step in using past labor policies to shed light on current developments is to highlight the recurrent dilemmas of managerial control over workers and work practices; a second is to construct a sequence of the crises and reforms in industrial relations strategies adopted by employers to cope with those dilemmas. The list, like any periodization, will be open to argument and revision. During the period covered by the last 100 years, I would include the following crises and strategic reforms: the years of experimentation with trade agreements around the turn of the century and the ensuing open shop drive between 1904 and 1914; the upheavals caused by labor insurgency and government intervention during World War I, followed by the 1920s "American plan"; the industrial conflict of the 1930s, ushering in the New Deal system; and the breakdown of the New Deal system in the 1980s and its replacement with a more "flexible," nonunion industrial relations regime. Why not include the 19th century (or earlier)? Before the late 1800s, (1) factory production, unionizing efforts, and a clear differentiation of labor and management were not yet the norm, and (2) employers' understanding of a "labor problem" was ill formed, at best. Making 1900 the cutoff thus provides for more "controlled" comparisons between periods and trains attention on what both observers today and social actors in 1900 would recognize as common dilemmas of management. Were this a self-standing study rather than an illustration of historical methods, these claims for temporal boundaries and periodization would need to be well documented.

Using turning points for periodization is a strategy shared by path dependency and reiterated problem solving, and in both cases it highlights a combination of continuities and contrasts across industrial relations regimes. Reiterated problem solving sheds additional light on contemporary antiunionism by building *multiple* switch points and paths into explanatory sequences and by constructing those sequences in different ways.

Consider the historical dynamic of successive battles over, and designs for, workplace control. Late 19th-century changes in production methods and the scale of industry both fueled unrest and pitted management prerogatives against restrictive craft customs. Indeed, it was in the late 19th century that a recognizably modern "labor problem," including a frank acknowledgment of separate and often opposed interests of labor and capital, became an object of widespread discussion and alarm among employers, politicians, and journalists (Rodgers 1974; Wiebe 1967). We are justified in treating the period as an important switch point not only because of this animated debate but also because employers advanced different diagnoses and championed different solutions for the problem. For example, trade agreements with provisions to regularize relations with unions and subject workplace conflicts to the rule of law were proposed as an alternative to open shops (Barnett 1912; Ramirez 1978). The logic of both path dependency and reiterated problem solving would then have us investigate the historically specific conditions that led most U.S. employers to opt for the open shop. These conditions might include contingent events. A machinists' strike in 1901, for example, caused employers to lose faith in unions as reliable partners in regulating the trade. Nudging industrial relations in the same direction were labor and product markets that gave employers an interest in purging unions, and state policies that gave them the means to do so (Haydu 1988).

Reiterated problem solving and path dependency would also point to mechanisms that sustained this open shop outcome over time, thus establishing a new industrial relations regime. Path dependency would likely highlight sunk costs in open shop management, including the development of generously staffed personnel departments, pressures on individual firms to conform to employer associations' practices, and the ways that open shop techniques became embedded in business school curricula and the profession of personnel management. These are important influences, but they are not the only influences. Putting problems and actors rather than outcomes at center stage reveals additional factors of interest. First, this shift in the narrative focus calls attention to the development of new ideological lenses through which employers viewed the labor problem. One example is the growing conviction that union recognition and an employer's right to manage could never be reconciled. Second, emphasizing problems and actors makes the analysis more sensitive to the ways that short-term accommodations created the tools that other problem solvers would wield in later periods. Open shops did more than deny unionists a collective voice. They also developed unilateral mechanisms for workplace regulation to take the place of craft traditions. Finally, the shift in focus brings into view the less obvious virtues of the open shop for dealing with employers' labor troubles. Insulating workplace labor relations from

unions had the additional advantage of discouraging ties between employees and community-wide labor institutions capable of mobilizing broader identities and action. Open shops in effect balkanized industrial conflict (Kimeldorf and Stepan-Norris 1992; Cornfield 1991; Haydu 1998). These are significant differences in emphasis. They do not, however, suggest sharp differences between path-dependent and problem-solving approaches in the underlying *description* of historical switch points and subsequent trajectories.

The decisive differences between these methodological strategies appear when considering how outcomes around the turn of the century influenced *later* turning points and choices and thus etched a larger trajectory. A model of reiterated problem solving is better suited both for tracing multiple registers of causation and for showing how outcomes not only set new paths but also fomented new crises. Early 20th-century outcomes led to current practices in part by consolidating an enduring language of management prerogatives, with its corresponding stigmatization of unions as illegitimate outsiders and usurpers of employers' rights. Some of the causal chains that link past and present industrial relations practices wind through considerably more devious routes. The victory of open shops after 1900 set in motion a differently paced historical trajectory of management reform. Against the backdrop of open shop orthodoxy, World War I precipitated a new crisis as unions, empowered by economic boom and government intervention, demanded recognition and supported more broadly based worker mobilization.

Path dependency would emphasize how these new forces—most of them external to the workplace—briefly put industrial relations on the New Deal track. This is, indeed, the most common view of wartime labor relations. According to this interpretation, worker mobilization and a more sympathetic and activist state led to government recognition of basic union rights and, in some industries, to a precocious capital-labor accord (Dubofsky 1994; Fraser 1983). Interpreting these developments as instances of reiterated problem solving affords a different view. The wartime crisis may have been sparked by exogenous forces, but it occurred along fault lines laid down earlier. And wartime problem solving was not a premature New Deal, aborted by the armistice. Instead, it updated the prewar status quo and set the agenda for subsequent reforms. Employers had to devise means to make management authority more constitutional and to grant representation rights to mobilized workers. At the same time, they had to erect new barricades between workplace labor relations and community-wide union institutions. As one key strategist of the open shop movement explained, labor organization could be appropriate for modern industry if it were based on “factory solidarity as opposed to class solidarity” and kept open to union and nonunion employees alike. “Intra-factory

organization of employees produces greater loyalty and solidarity between the management and the employees and thereby makes the men less susceptible to the appeal of militancy" (Walter Merritt 1919a, p. 1627; 1919b, p. 1706). The square deal and employee representation plans—including company unions—met these requirements (Jacoby 1985; Haydu 1997). Explaining the development of this "American plan" in the 1920s thus demands attention to a linked sequence of problems and resolutions. The unique dilemmas of wartime mobilization were critically important, but these dilemmas themselves had been structured by prior open-shop strategies. In turn, the postwar outcomes became part of the repertoire of personnel techniques passed along to later managers. Focusing on problem-solving action rather than on paths reveals how differently paced developments come together in a single turning point and how new paths are themselves charted on the basis of ideological lenses, strategic tools, and pressing problems inherited from prior crises.

There is still more to the causal lineage of contemporary antiunionism. The 1920s solution for workplace control and labor insurgency proved no more enduring than its prewar predecessor. By the late 1930s, the American plan yielded to the New Deal in mass production industries. Collective bargaining under the National Labor Relations Act, most commentators agree, set a new direction in U.S. industrial relations, whether this new direction is labeled the "institutionalization of industrial conflict" or the "capital-labor accord" (Tomlins 1985; Fraser and Gerstle 1989). And the usual explanations for this sharp turn are also consistent with path dependency: the exogenous factors of state intervention and economic collapse finally changed the course of American industrial relations. The Great Depression spurred some employers to begin mulling over the potential virtues of unions for achieving economic stability. Mass mobilization by industrial workers helped them make up their minds (Gordon 1994). More important, new political alignments transformed the state from a reliable business ally into a defender of minimal labor rights (Brody 1980).

Rethinking the New Deal system as part of a longer sequence of problem solving brings to light continuities missing from conventional accounts. For one thing, the 1920s industrial relations regime itself contributed to economic breakdown. A system that insulated labor relations from external influences, however useful for fragmenting industrial conflict, could make no contribution to macroeconomic regulation. The exclusion of mass production workers from industrial governance by craft unions and open shops in the 1920s was also fateful. It ensured that worker mobilization in the 1930s would take the form of industrial unionism and strong labor support for New Deal intervention. The crisis of the 1930s, then, was in part the product of prior problem solving by employers. When

employers finally embraced “workplace contractualism” as their best hope for restabilizing industrial relations, they continued yet another trajectory set earlier. During the 1920s, many larger employers had developed administrative mechanisms and professional staffs to dispense justice and organize representation unilaterally. Under the New Deal, they turned these investments to the purposes of collective bargaining and contract administration, so that even unionized labor relations remained centered on the individual plant or firm. Past history, moreover, gave unionists good reason to assume the worst of management and to rely on contractual protection as the best defense against abuses of authority (Brody 1993; Amberg 1991). Here again, the logic of reiterated problem solving reveals what path dependency obscures: the interweaving of multiple causal sequences.

It is in light of these prior sequences of problem solving, finally, that we can illuminate contemporary developments. Invoking globalization to explain the quest for union-free labor relations obviously misses the historical roots of employer strategies. But even path dependency’s use of the past, in which responses to external shocks either follow prior tracks or jump to new ones, obscures key historical legacies. Shifting attention to reiterated problem solving reveals how current employer policies reflect a more complex inheritance. It does so, first, by identifying the *multiple* paths that converge in today’s business practices. Pre–World War I battles over technological change made unions appear to be the implacable enemies of management control rather than, as in many European cases, unavoidable allies for codifying work rules and discipline. This early 20th-century perception of unions persists with only minor modifications today. Another inheritance followed a different path. Having repudiated collective bargaining as a way to handle labor relations, employers developed alternative, in-house mechanisms for regulating the workplace, including internal labor markets and welfare capitalism. These techniques have their own history. Over time, they have been retooled by the dynamics of professionalism as well as the needs of capital (Jacoby 1997). But this historical tributary converges with antiunion ideology today, as businesses deploy an inherited repertoire of human resources management to solve current problems.

It is not just employers’ responses that have multiple historical roots—so do the challenges they face. Path-dependent accounts of restructuring highlight the exogenous impact of globalization. Reiterated problem solving identifies how contemporary troubles developed out of employers’ own past strategies. More competitive and fragmented markets, after all, do not necessarily threaten unionized labor relations. They incite a crisis, and thus a turning point, only because of the specific character of old regimes. Most interpretations of contemporary antiunionism miss the fact

that the basic structure of New Deal labor relations (much of it inherited, in turn, from the 1920s) guarantees that the current quest for flexibility will pit management against unions. Unlike European systems of industry-wide bargaining, plant-level contracts regulate working conditions and job practices in detail, leaving factory managers little leeway for improvisation (Sisson 1991). This legalistic and exacting rule of law (a legacy both of unionists' self-protection in the 1930s and of management's own square deal a decade earlier) also runs counter to flexible and informal approaches to decision making and conflict resolution (Edwards 1993).

Employers do, of course, face unprecedented challenges today. But an explanation for their antiunion response cannot rely wholly on unique features of the present era, as an individualizing comparison would call for. Nor is it satisfactory to treat antiunionism, as Lipset would do, as a resurfacing of a causal constant—America's deeply rooted individualism. Finally, in contrast to path dependency, explanations of contemporary industrial relations must be multilayered and actor centered. Earlier turning points explain why employers define current problems in terms of the need for "flexibility"; many of today's human relations techniques were pioneered in earlier epochs of welfare capitalism; and contradictions between the New Deal system and globalization underlie the crisis that faced employers in the 1980s.

In summary, an identifiable series of past responses to union challenges left a rich causal legacy for antiunionism today. These past sequences both fuel and shape the current labor crisis, set the constraints that contemporary employers seek to overcome, make unions appear the main obstacles to progress, and bequeath models for the nonunion workplace that management can adapt to new conditions. By focusing attention on recurrent crises, problem-solving actors, and multiple registers of causality, the reiterated problem-solving approach provides more concrete guidance in how to construct these explanatory sequences than can be gleaned from most advocates of narrative. For these same reasons, problem solving is better suited than path dependency for putting the multiple legacies of the past to use in explaining antiunionism today.

CONCLUSION

Most historical sociologists recognize the value of dividing time into periods, each with characteristics distinguishing it from others. What those characteristics are, and thus where to draw the line between one period and another, are frequently disputed. But if the value of periodization of *some* kind is conceded, then certain methodological implications follow when information from one period is deployed to help interpret another.

Two ideal-typical strategies for comparing societies are replicated in the ways historical sociologists use the past. One searches for causal regularities across periods, and the other juxtaposes periods in order to enrich our understanding of the historical individuality of each. I have argued that the familiar weaknesses of those strategies reappear when comparing times instead of places, including the characteristic excesses of explanations without histories and histories without explanation. These comparative approaches also share certain limitations—and miss certain opportunities—that are peculiar to the comparison of times. The most important of these constraints is the inability to explain sequences of events over time.

Methods that put sequences of events at the center of analysis, including narrative and path dependency, strike a better balance between historically insensitive causal generalization and idiographic historicism. They are also better able to represent the ways that Henry Ford's "one damn thing after another" is actually historical causation working through time. Connecting events across periods in terms of sequenced problem solving has three additional virtues that distinguish it from narrative and path dependency. First, it provides a better sense than narrative accounts of the mechanisms that conduct causal influence over time. Second, it captures the creative as well as the constraining role of the past at each turning point, as path-dependent models do not. Inherited solutions limit options, but they also precipitate and shape later crises, along with the choices social actors deem viable. Third, it accomplishes these methodological goals in part by making social actors the historical pivots that link "cases" of problem solving. Reconstructing the problem-solvers' understandings and choices—how *they* make use of the past—enables us to account for trajectories across multiple periods.

Most sociologists have come to accept that a single method does not fill all empirical needs. Reiterated problem solving is no exception, even for multiperiod historical studies. Where applicable, however, it is a valuable device for organizing events into explanatory sequences and trajectories across multiple time periods, while still respecting the historical integrity of each of them.

REFERENCES

- Abbott, Andrew 1983 "Sequences of Social Events: Concepts and Methods for the Analysis of Order in Social Processes." *Historical Methods* 16 (4): 129–47.
- 1992a "From Causes to Events: Notes on Narrative Positivism." *Sociological Methods and Research* 20 (4): 428–55
- 1992b "What Do Cases Do? Some Notes on Activity in Sociological Analysis." Pp. 53–82 in *What Is a Case? Exploring the Foundations of Social Inquiry*, edited by Charles Ragin and Howard Becker. Cambridge. Cambridge University Press.

- . 1997. "On the Concept of Turning Point." *Comparative Social Research* 16: 85–105.
- Abbott, Andrew, and Alexandra Hrycak. 1990. "Measuring Resemblance in Sequence Data: An Optimal Matching Analysis of Musicians' Careers." *American Journal of Sociology* 96 (1): 144–85.
- Abrams, Philip. 1982. *Historical Sociology*. Ithaca, N.Y.: Cornell University Press.
- Amberg, Stephen. 1991. "Democratic Producerism: Enlisting American Politics for Workplace Flexibility." *Economy and Society* 20 (1): 57–78.
- Aminzade, Ronald. 1992. "Historical Sociology and Time." *Sociological Methods and Research* 20 (4): 456–80.
- Appelbaum, Eileen, and Rosemary Batt. 1994. *The New American Workplace: Transforming Work Systems in the United States*. Ithaca, N.Y.: ILR Press.
- Arthur, W. Brian. 1988. "Self-Reinforcing Mechanisms in Economics." Pp. 9–31 in *The Economy as an Evolving Complex System*, edited by Philip W. Anderson, Kenneth J. Arrow, and David Pines. Redwood City, Calif.: Addison-Wesley.
- Barnett, George. 1912. "National and District Systems of Collective Bargaining in the United States." *Quarterly Journal of Economics* 26 (May): 425–43.
- Bendix, Reinhard. 1970. *Embattled Reason: Essays on Social Knowledge*. New York: Oxford University Press.
- . (1956) 1974. *Work and Authority in Industry: Ideologies of Management in the Course of Industrialization*. Berkeley and Los Angeles: University of California Press.
- . 1978. *Kings or People: Power and the Mandate to Rule*. Berkeley and Los Angeles: University of California Press.
- Brody, David. 1980. *Workers in Industrial America: Essays in the 20th Century Struggle*. New York: Oxford University Press.
- . 1993. "Workplace Contractualism in Comparative Perspective." Pp. 176–205 in *Industrial Democracy in America: The Ambiguous Promise*, edited by Nelson Lichtenstein and Howell John Harris. Washington, D.C.: Woodrow Wilson Center Press; New York: Cambridge University Press.
- Brustein, William. 1996. *The Logic of Evil: The Social Origins of the Nazi Party, 1925–1933*. New Haven, Conn.: Yale University Press.
- Burawoy, Michael. 1979. *Manufacturing Consent: Changes in the Labor Process under Monopoly Capitalism*. Chicago: University of Chicago Press.
- . 1989. "Two Methods in Search of Science: Skocpol versus Trotsky." *Theory and Society* 18 (6): 759–805.
- Campbell, Donald T., and H. Laurence Ross. 1968. "The Connecticut Crackdown on Speeding: Time-Series Data in Quasi-Experimental Analysis." *Law and Society Review* 3 (1): 33–53.
- Collier, David, and James Mahoney. 1996. "Insights and Pitfalls: Selection Bias in Qualitative Research." *World Politics* 49 (October): 56–91.
- Cornfield, Daniel B. 1991. "The U.S. Labor Movement: Its Development and Impact on Social Inequality and Politics." *Annual Review of Sociology* 17: 27–49.
- Craypo, Charles, and Bruce Nissen, eds. 1993. *Grand Designs: The Impact of Corporate Strategies on Workers, Unions, and Communities*. Ithaca, N.Y.: ILR Press.
- David, Paul A. 1986. "Understanding the Economics of QWERTY: The Necessity of History." Pp. 30–49 in *Economic History and the Modern Economist*, edited by William N. Parker. Oxford: Basil Blackwell.
- Dubofsky, Melvyn. 1994. *The State and Labor in Modern America*. Chapel Hill: University of North Carolina Press.
- Edwards, Richard. 1993. *Rights at Work: Employment Relations in the Post-Union Era*. Washington, D.C.: Brookings Institution.
- Fraser, Steve. 1983. "Dress Rehearsal for the New Deal. Shop-Floor Insurgents, Political Elites, and Industrial Democracy in the Amalgamated Clothing Workers." Pp.

- 212–55 in *Working-Class America: Essays on Labor, Community, and American Society*, edited by Michael H. Frisch and Daniel J. Walkowitz. Urbana: University of Illinois Press.
- Fraser, Steve, and Gary Gerstle, eds. 1989. *The Rise and Fall of the New Deal Order, 1930–1980*. Princeton, N.J.: Princeton University Press.
- Fredrickson, George M. 1988. *The Arrogance of Race: Historical Perspectives on Slavery, Racism, and Social Inequality*. Middletown, Conn.: Wesleyan University Press.
- Gans, Herbert J. 1995. *The War against the Poor: The Underclass and Antipoverty Policy*. New York: Basic Books.
- Goldstone, Jack. 1991. *Revolution and Rebellion in the Early Modern World*. Berkeley and Los Angeles: University of California Press.
- . 1997. "Methodological Issues in Comparative Macrosociology." *Comparative Social Research* 16: 107–20.
- Goldthorpe, John H. 1997. "Current Issues in Comparative Macro Sociology: A Debate on Methodological Issues." *Comparative Social Research* 16:1–26, 121–32.
- Gordon, Colin. 1994. *New Deals: Business, Labor and Politics in America, 1920–1935*. Cambridge: Cambridge University Press.
- Gotham, Kevin Fox, and William G. Staples. 1996. "Narrative Analysis and the New Historical Sociology." *Sociological Quarterly* 37 (3): 481–501.
- Gould, Roger V. 1995. *Insurgent Identities: Class, Community, and Protest in Paris from 1848 to the Commune*. Chicago: University of Chicago Press.
- Griffin, Larry J. 1992. "Temporality, Events, and Explanation in Historical Sociology." *Sociological Methods and Research* 20 (4): 403–27.
- . 1993. "Narrative, Event-Structure Analysis, and Causal Interpretation in Historical Sociology." *American Journal of Sociology* 98 (5): 1094–133.
- Haimson, Leopold, and Charles Tilly, eds. 1989. *Strikes, Wars, and Revolutions in an International Perspective: Strike Waves in the Late Nineteenth and Early Twentieth Centuries*. Cambridge: Cambridge University Press.
- Hall, Peter M. 1995. "The Consequences of Qualitative Analysis for Sociological Theory: Beyond the Microlevel." *Sociological Quarterly* 36 (2): 397–423.
- Haydu, Jeffrey. 1988. "Employers, Unions, and 'American Exceptionalism.'" *International Review of Social History* 33 (1): 25–41.
- . 1997. *Making American Industry Safe for Democracy: Comparative Perspectives on the State and Employee Representation in the Era of World War I*. Urbana: University of Illinois Press.
- . 1998. "Workplace Governance, Class Formation, and the New Institutionalism." *Mobilisation* 3 (1): 69–88.
- Hirschhorn, Larry. 1997. *Reworking Authority: Leading and Following in the Post-Modern Organisation*. Cambridge, Mass.: MIT Press.
- Isaac, Larry W. 1997. "Transforming Localities: Reflections on Time, Causality, and Narrative in Contemporary Historical Sociology." *Historical Methods* 30 (1): 4–12.
- Isaac, Larry W., and Larry J. Griffin. 1989. "Ahistoricism in Time-Series Analysis of Historical Process: Critique, Redirection, and Illustrations from U.S. Labor History." *American Sociological Review* 54 (6): 873–90.
- Jacoby, Sanford M. 1985. *Employing Bureaucracy: Managers, Unions, and the Transformation of Work in American Industry, 1900–1945*. New York: Columbia University Press.
- . 1997. *Modern Manors: Welfare Capitalism since the New Deal*. Princeton, N.J.: Princeton University Press.
- Katz, Michael. 1996. *In the Shadow of the Poorhouse: A Social History of Welfare in America*, rev. ed. New York: Basic Books.
- Kimeldorf, Howard. 1988. *Reds or Rackets? The Making of Radical and Conservative Unions on the Waterfront*. Berkeley and Los Angeles: University of California Press.

- Kimeldorf, Howard, and Judith Stepan-Norris. 1992. "Historical Studies of Labor Movements in the United States." *Annual Review of Sociology* 18:495-517.
- Kiser, Edgar, and Michael Hechter. 1991. "The Role of General Theory in Comparative-Historical Sociology." *American Journal of Sociology* 97 (1): 1-30.
- Kochan, Thomas A., Harry C. Katz, and Robert B. McKersie. 1986. *The Transformation of American Industrial Relations*. New York: Basic Books.
- Lamont, Michelle. 1992. *Money, Morals, and Manners: The Culture of the French and the American Upper-Middle Class*. Chicago: University of Chicago Press.
- Liebersohn, Stanley. 1987. *Making It Count: The Improvement of Social Research and Theory*. Berkeley and Los Angeles: University of California Press.
- Lipset, Seymour Martin. 1977. "Why No Socialism in the United States?" Pp. 31-149 in *Sources of Contemporary Radicalism*, edited by S. Bialer and S. Sluzar. Boulder, Colo.: Westview Press.
- . 1986. "North American Labor Movements: A Comparative Perspective." Pp. 421-52 in *Unions in Transition: Entering the Second Century*, edited by Seymour Martin Lipset. San Francisco: Institute for Contemporary Studies.
- . 1991. "American Exceptionalism Reaffirmed." Pp. 1-45 in *Is America Different? A New Look at American Exceptionalism*, edited by Byron E. Shafer. Oxford: Clarendon Press.
- Maines, David R. 1993. "Narrative's Moment and Sociology's Phenomena: Toward a Narrative Sociology." *Sociological Quarterly* 34 (1): 17-38.
- Marshall, T. H. 1964. *Class, Citizenship, and Social Development*. Garden City, N.Y.: Doubleday.
- McDaniel, Tim. 1988. *Autocracy, Capitalism, and Revolution in Russia*. Berkeley and Los Angeles: University of California Press.
- . 1996. *The Agony of the Russian Idea*. Princeton, N.J.: Princeton University Press.
- Merritt, Walter. 1919a. *Iron Age*, June 16, 1627.
- . 1919b. *Iron Age*, June 26, 1706.
- Moody, Kim. 1997. *Workers in a Lean World: Unions in the International Economy*. London: Verso.
- North, Douglass. 1990. *Institutions, Institutional Change and Economic Performance*. Cambridge: Cambridge University Press.
- Orloff, Ann Shola. 1988. "The Political Origins of America's Belated Welfare State." Pp. 37-80 in *The Politics of Social Policy in the United States*, edited by Margaret Weir, Ann Shola Orloff, and Theda Skocpol. Princeton, N.J.: Princeton University Press.
- Piore, Michael J., and Charles F. Sabel. 1984. *The Second Industrial Divide: Possibilities for Prosperity*. New York: Basic Books.
- Przeworski, Adam, and Henry Teune. 1970. *The Logic of Comparative Social Inquiry*. New York: Wiley.
- Quadagno, Jill, and Stan J. Knapp. 1992. "Have Historical Sociologists Forsaken Theory? Thoughts on the History/Theory Relationship." *Sociological Methods and Research* 20 (4): 481-507.
- Ragin, Charles. 1987. *The Comparative Method: Moving beyond Qualitative and Quantitative Strategies*. Berkeley and Los Angeles: University of California Press.
- Ramirez, Bruno. 1978. *When Workers Fight: The Politics of Industrial Relations in the Progressive Era, 1898-1916*. Westport, Conn.: Greenwood Press.
- Rodgers, Daniel T. 1974. *The Work Ethic in Industrial America, 1850-1920*. Chicago: University of Chicago Press.
- Róna-Tas, Ákos. 1998. "Path-Dependence and Capital Theory Sociology of the Post-Communist Economic Transformation." *East European Politics and Societies* 12:107-31.

- Roy, William G. 1997. *Socialising Capital: The Rise of the Large Industrial Corporation in America*. Princeton, N.J.: Princeton University Press.
- Rueschemeyer, Dietrich, and John D. Stephens. 1997. "Comparing Historical Sequences—a Powerful Tool for Causal Analysis." *Comparative Social Research* 16: 55–72.
- Sewell, William H., Jr. 1980. *Work and Revolution in France: The Language of Labor from the Old Regime to 1848*. Cambridge: Cambridge University Press.
- . 1996. "Three Temporalities: Toward an Eventful Sociology." Pp. 245–80 in *The Historic Turn in the Human Sciences*, edited by Terrence J. McDonald. Ann Arbor: University of Michigan Press.
- Sisson, Keith. 1991. "Employers and the Structure of Collective Bargaining: Distinguishing Cause and Effect." Pp. 256–71 in *The Power to Manage? Employers and Industrial Relations in Comparative-Historical Perspective*, edited by Steven Tollday and Jonathan Zeitlin. London: Routledge.
- Skeels, Jack W. 1982. "The Economic and Organizational Basis of Early United States Strikes, 1900–1948." *Industrial and Labor Relations Review* 35 (4): 491–503.
- Skocpol, Theda. 1979. *States and Social Revolutions: A Comparative Analysis of France, Russia, and China*. Cambridge: Cambridge University Press.
- . 1992. *Protecting Soldiers and Mothers: The Political Origins of Social Policy in the United States*. Cambridge, Mass.: Harvard University Press.
- , ed. 1984. *Vision and Method in Historical Sociology*. Cambridge: Cambridge University Press.
- Somers, Margaret R. 1992. "Narrativity, Narrative Identity, and Social Action: Rethinking English Working-Class Formation." *Social Science History* 16 (4): 591–630.
- . 1996. "Where Is Sociology after the Historic Turn? Knowledge Cultures, Narrativity, and Historical Epistemologies." Pp. 53–89 in *The Historic Turn in the Human Sciences*, edited by Terrence J. McDonald. Ann Arbor: University of Michigan Press.
- Steinmetz, George. 1992. "Reflections on the Role of Social Narratives in Working-Class Formation: Narrative Theory in the Social Sciences." *Social Science History* 16 (3): 489–516.
- Stinchcombe, Arthur L. 1978. *Theoretical Methods in Social History*. New York: Academic Press.
- Thompson, E. P. 1963. *The Making of the English Working Class*. New York: Vintage Books.
- Tilly, Charles. 1984. *Big Structures, Large Processes, Huge Comparisons*. New York: Russell Sage.
- . 1997. "Means and Ends of Comparison in Macrosociology." *Comparative Social Research* 16:43–53.
- Tomlins, Christopher. 1985. *The State and the Unions: Labor Relations, Law, and the Organized Labor Movement in America, 1880–1960*. Cambridge: Cambridge University Press.
- Voss, Kim. 1993. *The Making of American Exceptionalism: The Knights of Labor and Class Formation in the Nineteenth Century*. Ithaca, N.Y.: Cornell University Press.
- Wiebe, Robert H. 1967. *The Search for Order, 1877–1920*. New York: Hill & Wang.
- Wilentz, Sean. 1984. *Chants Democratic: New York City and the Rise of the American Working Class, 1788–1850*. New York: Oxford University Press.
- . 1997. "The Past Is Not a 'Process.'" *New York Times*, April 20, E15.

Defended Neighborhoods, Integration, and Racially Motivated Crime¹

Donald P. Green, Dara Z. Strolovitch, and Janelle S. Wong
Yale University

This article investigates demographic and macroeconomic correlates of racially motivated antiminority crime in New York City (1987–95). Event count models indicate that crimes directed against Asians, Latinos, and blacks are most frequent in predominantly white areas, particularly those that had experienced an in-migration of minorities. No relationship is found between rates of racially motivated crime and macroeconomic conditions, such as the rate of unemployment among non-Hispanic whites; nor does there appear to be an interaction between economic conditions and in-migration of minorities. These findings seem to parallel ethnographic accounts of “defended” white urban neighborhoods. The article concludes by discussing the empirical implications of this theoretical perspective as applied to prejudice-based crime in other contexts.

INTRODUCTION

The study of racial prejudice has for decades occupied a central place in the social sciences, and yet certain basic research questions remain largely unexamined. One such question concerns the conditions under which racially motivated crimes occur. Is the frequency of racial harassment and violence higher in areas where minorities make up a small proportion of the population or where their numbers are on par with those of the dominant racial group? A related question concerns the link between demographic change and racially motivated crime. To what extent does minority victimization depend on the ways in which the proportions of different racial groups have changed over time? Finally, to what extent and in what

¹ The authors are grateful to the Bias Crime Unit of the New York Police Department (NYPD) for making their data available to us. This research was made possible by awards from the National Science Foundation (SBR935-7937) and the Institution for Social and Policy Studies at Yale University, neither of which bear responsibility for the findings presented here. The data used in this study and supplementary statistical analyses may be found at the Web site <http://pantheon.yale.edu/~gogreen/ajs repl.html>. Direct correspondence to Donald Green, Department of Political Science, Yale University, 124 Prospect Street, New Haven, Connecticut 06520-8301.

ways do economic circumstances affect the incidence of acts of intimidation directed at minorities?

This article seeks to address these questions by analyzing patterns of racially motivated crime in New York City and drawing parallels to other case studies. Notwithstanding dramatic growth in the number of laws mandating harsher penalties for crimes motivated by racial animus, the systematic analysis of racially motivated crime remains in its infancy, and our aim is to lay the groundwork for studies of greater geographic and historical scope. Our principal finding is that demographic change, not economic hardship or inequality, predicts racially motivated crime directed at minorities. Victimization of blacks, Asians, and Latinos tends to be higher in areas where whites have long been the predominant racial group, particularly when these areas experience significant growth in minority population. On the other hand, we turn up no relationship between unemployment rates and racially motivated crime. This finding suggests a proposition with far-reaching social and theoretical implications: racially motivated crime stems not from economic frustration but from an exclusionary impulse on the part of whites defending what they perceive to be their territory.

THEORIES ABOUT THE CONSEQUENCES OF RACIAL BALANCE

Efforts to collect data on the incidence of "bias crime," or illegal conduct motivated by animus toward the victim's race, religion, ethnicity, or sexual orientation, began in the 1980s, and only recently have scholars undertaken systematic analyses of bias-crime patterns (Comstock 1991; Garofalo 1991; Green, Glaser, and Rich 1998; Green and Rich 1998; Krueger and Pischke 1997). Because demographic change has not been the focal point of extant research, intuitions about the relationship between racial composition and racially motivated crime must be gleaned from cognate literatures concerning racial attitudes, housing segregation, and racial violence. These literatures, however, offer conflicting hypotheses about the conditions under which racially motivated violence, harassment, and intimidation are likely to arise.

On the one hand, hypotheses derived from realistic group conflict theory (Levine and Campbell 1972; Bobo 1988) suggest that attacks on minorities will be more frequent where their numbers are large. One line of argument, known as the power-threat hypothesis (Blalock 1967; Tolnay, Beck, and Massey 1989), holds that white intolerance is greatest where the size of the nonwhite population presents a challenge to the economic interests and dominant social and political position of whites. Consistent with this line of research are patterns of antiblack lynching in the prewar South (Tolnay and Beck 1995) and white support for ex-Klansman David Duke

in Louisiana's 1990 race for U.S. Senate (Giles and Buckner 1993), both of which were greatest in areas with high concentrations of blacks.

These findings parallel survey evidence concerning the "social distance" whites prefer to keep between themselves and racial minorities. A desire to guard their group interests, argues Bobo (1988), makes whites apprehensive about large numbers of minorities settling into their neighborhoods (Schuman, Steeh, and Bobo 1985). Whites profess a desire to live in an integrated setting but one in which they remain the predominant racial group (Farley et al. 1978; Clark 1991; Bobo, Schuman, and Steeh 1986).² If racially motivated crime reflects whites' stated attitudes regarding social distance, we should expect attacks to be more numerous in areas where minority populations have grown to the point at which white predominance is threatened.

Another variant of realistic group conflict theory suggests that interracial conflict intensifies as the minority population approaches parity with the white population, competing with it for jobs and housing, but declines as the minority population eclipses the size and power of the white population. This curvilinear pattern may be accentuated by the political dynamics set in motion as public officials and interest group leaders mobilize and, in so doing, polarize different racial constituencies (Horowitz 1985). Racially motivated crime, on this account, would represent a street-level manifestation of a broader political struggle. A curvilinear pattern, on the other hand, need not hinge on the notion that racial antagonism is inspired or orchestrated by elites. If interracial frictions are proportional to the amount of interracial contact in a given area, racially motivated crime should be higher in racially heterogeneous areas (Blau 1977). Although the evidence is mixed, there is some indication that this pattern holds for interracial crime, some of which may be racially motivated (Sampson 1984; South and Messner 1986).

A slightly different prediction seems to flow from the literature on white flight and so-called tipping points. Consistent with the realistic group conflict perspective, this research suggests that the trepidation of white home owners grows as racial minorities achieve more than a token presence in a region. No consensus exists on the question of how high the minority fraction must be before tipping occurs, but most studies tend to find that the exodus of whites accelerates after blacks constitute at least a quarter

² In 1987, when asked their preferred racial mix assuming they "could live in a neighborhood as good as the one [they] are living in now," 10% of white New Yorkers stated a preference for all-white neighborhoods, 20% for mostly white, 41% for half white/half black, 9% for a multiethnic neighborhood, with the remainder declining to state a preference (*New York Times* 1987).

of the population in a residential area (Galster 1990; Ottensmann and Gleeson 1992; Clark 1993). If one assumes, as many scholars do, that out-migration expresses the hostile racial sentiments of whites, one might also suppose that racially motivated crime would follow an analogous pattern, reaching its flashpoint sometime before minorities constitute half of the population and diminishing thereafter.¹

While some studies imply that the incidence of racially motivated crime will be greatest in neighborhoods where minorities comprise a substantial fraction of the population, others suggest that racially motivated incidents will be most frequent when minorities constitute a small share of the population. Acts of racial intimidation arguably reflect power differences between groups. Members of the dominant group may be emboldened to attack by the perception that law enforcement officials and the majority of those living in the neighborhood are unsympathetic to the victim group (Myrdal 1944; Levin and McDevitt 1993). By the same token, where minorities are few in number, perpetrators have less to fear by way of reprisal. We shall refer to this possibility as the power-differential hypothesis (Levine and Campbell 1972).

Working in tandem with power differences are social psychological mechanisms that encourage attacks against outsiders in predominantly white areas. Suttles (1972) and other ethnographers of white urban neighborhoods (Rieder 1985; De Sena 1990) argue that residents perceive that they and others of the same race living nearby share a common identity that hinges on the exclusion of racial or ethnic outgroups. This identity, as well as the social networks that work to safeguard a neighborhood's putative characteristics and ways of life, becomes salient in homogeneous white areas confronted with integration. Exclusion of minorities may obviously be accomplished by means other than harassment and intimidation, but widespread concern about stemming the tide of integration creates an environment in which recourse to more extreme behaviors is encouraged or at least tolerated (Rieder 1985). Although authors writing about defended neighborhoods often underscore the exaggerated fears of white residents, these studies differ markedly from the "symbolic politics"

¹ Within an urban setting, socioeconomic factors may enhance the possibility that racial intimidation will become more frequent as the minority population grows. The literature concerning white reactions to policy interventions such as court-ordered school desegregation points out that affluent whites, who tend to express the most tolerant attitudes concerning blacks, also possess the economic means to relocate to avoid what are perceived to be the adverse consequences of integration (Smith 1982; Bobo and Licari 1989; Carmines and Stimson 1989). The exit of more affluent whites changes both the distribution of opinion within the neighborhood and the composition of its most influential and politically active stratum.

literature, which reports no connection between racial context and prejudiced attitudes (Sears and Citrin 1997). The literature on defended neighborhoods may be interpreted as a rapprochement between symbolic and realistic perspectives: while emphasizing the importance of nonmaterial values (such as preserving a way of life), ethnographers trace conflict over neighborhood territory to the onset of racial integration.

Like the power-differential hypothesis, studies of defended neighborhoods suggest that higher rates of racially motivated crime will occur in areas where whites enjoy numerical superiority. The mechanisms are somewhat different, however. The power-differential interpretation emphasizes the capacity of minorities to protect themselves as their numbers grow. From the vantage point of defended neighborhood studies, growing minority populations undermine the preexisting social networks that both foster whites' sentimental attachment to a racially homogeneous image of the community (Suttles 1972, p. 35) and facilitate acts of hostility against outsiders by the most belligerent community members (Reider 1985). As integration moves forward, some whites adapt to the notion that nonwhites are an expected feature of neighborhood life (Alba et al. 1994; Kinder and Mendelberg 1995), while those most hostile to minority encroachment quit the neighborhood after initial efforts to drive away newcomers fail. In the latter case, the apparent improvement in race relations as neighborhoods become integrated reflects a progression from white fight to white flight.

The power-differential and defended neighborhood perspectives, however, offer divergent predictions about the consequences of demographic *change*. From a power-differential perspective, an influx of minorities signifies a diminution of white power, whereas from a defended neighborhood vantage point, it represents the catalyst for action among those who seek to preserve racial homogeneity. Thus, one key feature of the defended neighborhood thesis is the interaction it posits between white homogeneity and the rate at which that homogeneity is being eroded. A similar kind of argument might be derived from a realistic group conflict perspective (Olzak, Shanahan, and West 1994, p. 201) on the grounds that whites seek to defend the material resources associated with their racially homogeneous neighborhoods. Unlike the defended neighborhood thesis, however, which predicts acts of racial intimidation will diminish once significant numbers of minorities establish themselves within a neighborhood, the resource conflict perspective anticipates a heightening of tensions as minority and majority populations reach parity.

This array of competing predictions is summarized in figure 1. The power-threat hypothesis suggests that racially motivated crime grows monotonically with the nonwhite proportion of the population. The power-differential hypothesis predicts the opposite, that such crimes will

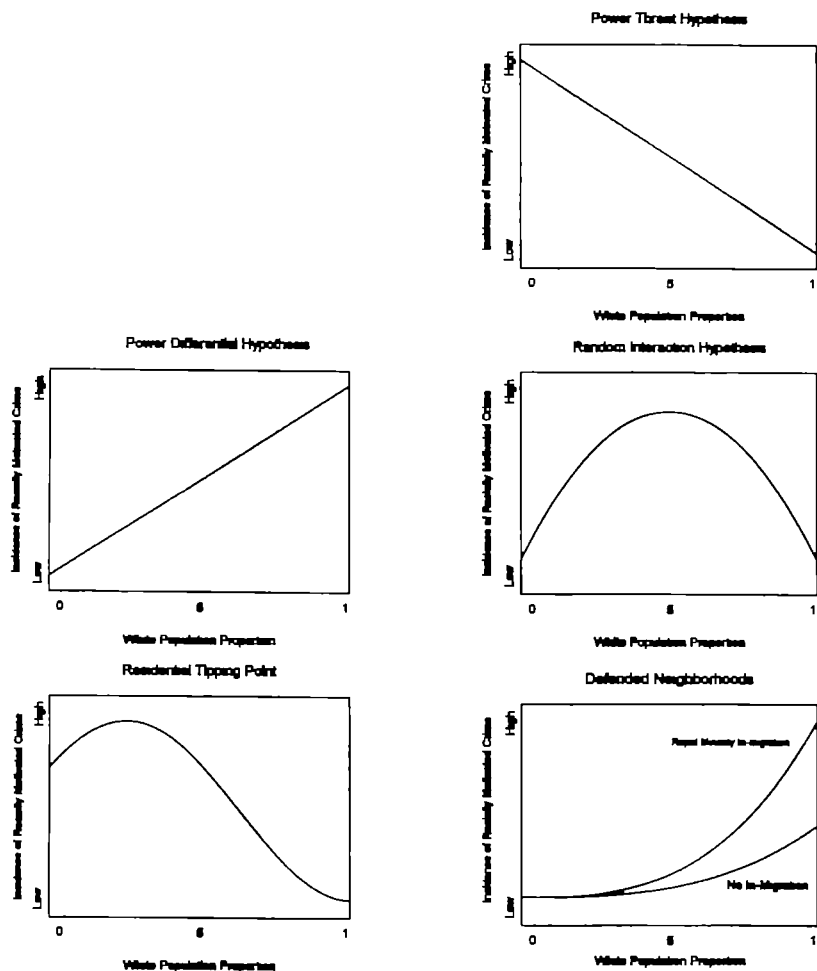


FIG. 1.—Illustration of competing hypotheses concerning racially motivated crime.

diminish as nonwhites gain in numbers and, by extension, in the capacity to defend themselves. Inverted U-shaped predictions flow from the random-interaction model, the structure of white preferences concerning levels of residential integration, and the tipping-point model. What we shall call for convenience the “defended neighborhood model,” illustrated at the bottom of figure 1, posits an accelerating relationship whereby racially motivated crimes are most frequent in areas where whites have long predominated, particularly those areas in which the minority population has begun to increase.

To this point, we have restricted our attention to hypotheses linking racially motivated crime to racial composition and have set aside the most prominent set of theoretical arguments, those linking acts of racial animus to economic conditions. As noted above, realistic group conflict theory traces intergroup hostilities to competition for scarce resources, a struggle that becomes more acute during periods of macroeconomic contraction (Tolnay and Beck 1995; Olzak 1992). We should therefore expect to see higher rates of antiminority crime in areas where whites are most vulnerable economically. When white unemployment is high, for example, unlawful incidents directed against minorities should become more common (Pinderhughes 1993). Another plausible variant of this hypothesis is that racially motivated crime surges where economically vulnerable whites confront an in-migration of minorities. It is argued, for example, that widespread unemployment precipitated the wave of anti-immigrant violence in the former East Germany following reunification (Willems 1995, p. 173) and that racial violence in the United States during the early 20th century grew out of job competition between blacks and whites during periods of economic uncertainty and rapid rates of immigration (Olzak 1989, 1992). In other work, we have examined in detail monthly time-series of racially motivated crime in New York city and found no connection whatsoever to fluctuations in unemployment rates (Green et al. 1998). In the cross-sectional analysis that follows, we look for possible interactions between economic strain and changing racial composition.

OVERVIEW OF RACIALLY MOTIVATED CRIME AND DEMOGRAPHIC CHANGE IN NEW YORK CITY

New York typifies the emerging multicultural composition of American cities.⁴ The number of multiethnic neighborhoods is increasing, and all-white neighborhoods are becoming much less common. In 1980, non-Latino whites comprised 52% of New York city's population; a decade later, this figure had ebbed to 43% (New York Department of City Planning 1992). As the white population of the city diminished, a concomitant

⁴ New York City's demographic changes are similar to those in other U.S. cities. For instance, whites comprise 40% of the population of New York, 38% of Chicago, and 36% of Los Angeles. Over one-quarter of New York City's population is black (28%), and the same is true of Boston (26%) and Pittsburgh (26%) (U.S. Bureau of the Census 1994). Large cities across the United States, such as Los Angeles, Chicago, and Miami, have experienced a growth in immigrant population (Denton and Massey 1991). Immigrants make up 28% of the population in New York, 20% of the population in Boston, and 34% of the population in San Francisco. Indices of social interaction and segregation for other large cities are comparable to those calculated for New York City (Massey and Denton 1987).

decline occurred in the number of predominantly white neighborhoods. In 1980, 15 of 59 so-called community districts⁵—groupings of census blocks containing approximately 130,000 people used by the Department of Planning—were at least 80% non-Hispanic white. Just seven such community districts remained by 1990. Similar transformations occurred in cities throughout the United States during this period (Thernstrom and Thernstrom 1997, chap. 8).

Although New York City as a whole is ethnically diverse, aggregate figures conceal its racial and ethnic enclaves. For example, in 1990, Brooklyn Community District 17, which includes the neighborhoods of Flatbush, Rugby, and Farragut, was 3.5% white and 88% black. In contrast, nearby Brooklyn Community District 15, Sheepshead Bay, was mostly white in 1990 (84%), with a population that was 7% Asian, 6% Latino, and 2% black. Similarly, Manhattan Community District 8, the Upper East Side, was 87.2% white in 1990. Across town, Manhattan Community District 9 (Manhattanville, Hamilton Heights, and Morningside Heights) was 19.5% white, 39.1% black, 36.1% Latino, and 4.5% Asian. The majority of the population in Staten Island is white, but even there differences exist across community districts. For instance, in 1990, in the northernmost community district, whites made up 65.2% of the population, and blacks were the next largest group (18%). In Staten Island's other two districts, whites accounted for over 85% of the population in 1990.

The form and rate of demographic change occurring in each community district has also varied. Brooklyn Community District 4, Bushwick, was 26.4% black and 56.4% Latino in 1980. Ten years later, the black population had declined slightly to 24.9%, while the Latino population had increased by nearly 10 percentage points to account for 65% of the total population. In Queens Community District 9 (Woodhaven, Richmond Hill, and Kew Gardens), the Asian population was less than 3% in 1980; by 1990, Asians made up nearly 10% of all residents. In Bronx Community District 7 (Fordham, Bedford Park, and Norwood), Latinos were 33.4%

⁵ The community district—a unit used by city agencies to represent, in rough fashion, communities in the city—serves as the unit of analysis throughout. Brooklyn, e.g., contains 18 community districts; Queens, 10. The Department of City Planning describes 12 community districts in the borough of the Bronx; however, as noted above, we have combined four of the Bronx districts with the smallest population levels into two districts in order to keep population levels as uniform as possible across districts and simplify our analysis. We did perform the analysis without combining the four districts into two and found no difference in results. We have included only upper Manhattan in our analysis (Manhattan Community Districts 7–12) because bias-crime data for the rest of Manhattan were not available at the community district level. Staten Island, the smallest New York City borough, contains three community districts.

of the population in 1980. In 1990, Latinos had become the majority group in the community district (50.6%).

Like most large cities, New York harbors both a diverse population and an undercurrent of intolerance. In 1986, a group of 20 whites attacked three black men in Howard Beach, in the borough of Brooklyn, stunning the city and the nation. Since that time, Brooklyn alone has witnessed nearly 300 racially motivated crimes directed against blacks, 84 directed against Latinos, and 78 directed against Asians. New York City created an investigative unit in 1981 within the New York Police Department (NYPD) that specifically deals with crimes motivated by prejudice toward the victim's perceived race, religion, ethnicity, or sexual orientation. Our data were assembled from individual incident reports to this Bias Crime Unit for the years 1987 to 1995. Reports for previous years are not available to us and, at any rate, cover a period before the procedures and staff size of the Bias Unit stabilized. Although these reports occasionally give the precise location of an incident, most reports list only the precinct in which the incident occurred. We aggregated the crime data from precincts to the community district level because community districts are more homogeneous in terms of population size.

All studies of bias crime inevitably confront thorny problems of measurement and conceptualization (Bowling 1993). Our data consist of incidents that were classified and investigated as bias crimes. Note that throughout we analyze the number of discrete crime incidents rather than the number of victims. Classification of incidents is based on a Bias Unit investigator's assessment of whether the unlawful conduct was motivated by racial animus on the part of the perpetrator. This assessment generally hinges on the language used by the perpetrator (e.g., racial slurs), the context in which the crime takes place (e.g., an ongoing pattern of harassment), and the absence of other motives (e.g., material gain). As shown in table 1, racially motivated crimes classified by the Bias Crime Unit comprise an array of illegal activity, from vandalism to harassment to murder. The bulk of the incidents directed at racial minorities, however, fall into the categories assault and harassment.⁶ Some illustrative examples culled from summary reports convey the nature of these incidents:

A 33-year-old black woman was at a bus stop with her son. A group of white men in a van drove by, sprayed her with a foam substance, and called her a "nigger." (Aggravated harassment, Precinct 122, February 1993)

A 48-year-old black male bus driver pulled over to use the phone. A vehicle pulled in front of the bus. A white male got out of the car and said, "Mother fuckin nigger get the fuck out of college point or I'll blow your

⁶ By contrast, 49% of the 1,296 anti-Semitic incidents over this period were classified as criminal mischief.

TABLE 1
RACE/ETHNICITY OF VICTIM AND TYPE OF CRIME REPORTED

Type of Crime	Black Victim (%)	Latino Victim (%)	Asian Victim (%)
Murder4	.8	0
Assault	26.4	41.2	39.8
Menacing	4.4	3.5	2.2
Harassment/aggravated harassment	51.3	42.4	34.4
Criminal mischief	12.0	5.5	16.1
Arson8	.8	.6
Robbery	1.5	3.1	3.9
Burglary	8	8	1.7
Unknown/other	2.4	2.0	2.2

NOTE.—*N* = 1,002 for black victims, 255 for Latino victims, and 180 for Asian victims

fuckin brains out." The bus driver was followed by the car as he drove out of the area. (Aggravated harassment, Precinct 109, April 1994)

A 19-year-old Asian male was on the train platform at the 59th and 4th Avenue station and was approached by a group of teenagers who asked him, "What you looking at Chino?" and "Hey Chinese." He was assaulted by the group, jumped to the tracks to escape, was followed by one of the teenagers, but eventually escaped. (Assault, Precinct 72, April 1994)

A 29-year-old Latino man was coming out of Pelham Bay Park when two to four white males jumped him. They punched him, called him a "stupid Puerto Rican," and told him to move off the block. They also told him that next time they would kill him. (Assault, Precinct 45, July 1994)

In each of the incidents described, the race of the perpetrator(s) was reported to the police. In many cases, however, particularly those involving threatening notes and vandalism, the perpetrator's race is not known or reported. As shown in table 2, lack of information about the perpetrator is the rule rather than exception. But as noted in table 3, there is a remarkable correspondence between (and only between) racially motivated crimes perpetrated by whites and those committed by an unidentified perpetrator. For example, the correlation between the number of antiblack crimes committed by whites and the number committed by unknown offenders is .82 across the 51 community districts in our sample. By contrast, the correlation between the number of antiblack incidents committed by Latinos correlates at $-.05$ with antiblack hate crimes committed by unidentified perpetrators. In supplementary analyses of the models described below (Green 1998), we find that one cannot reject the null hypothesis that the parameters that generate incidents by white offenders also generate incidents by unknown offenders. The parameters themselves look very similar after the data are disaggregated by perpetrator, although

TABLE 2

PERCENTAGE OF RACIALLY MOTIVATED CRIMES COMMITTED AGAINST EACH GROUP
BY PERPETRATOR RACE/ETHNICITY

Race/Ethnicity of Perpetrator	Black Victim (%)	Latino Victim (%)	Asian Victim (%)
Latino	2.6	.8	8.3
White	29.9	31.4	18.9
Black	1.9	11.8	15.6
Asian	1	0	1.1
Latino and black	1	1.2	6
White and Latino or black8	.8	.6
Other	4	0	0
Unknown	64.3	54.1	55.0

NOTE.— $N = 1,002$ for black victims, 255 for Latino victims, and 180 for Asian victims.

the smaller number of incidents in each perpetrator category makes for greater sampling variability. In order to maximize the precision with which we estimate our models, we focus our attention on all incidents involving white or unknown perpetrators.

The question remains, however, whether the incident reports themselves are reliable indicators of hate crime. The information gathered by the Bias Unit is doubtless incomplete. Victims may be reluctant to report incidents because they fear that others will view their racial or ethnic identity in a negative light or that they will become targets for further

TABLE 3

CORRELATIONS BETWEEN RACIALLY MOTIVATED CRIMES WITH KNOWN
AND UNKNOWN PERPETRATORS

Race of Known Perpetrators	Incidence of Racially Motivated Crime Committed by Unknown Perpetrators		
	Black Victims	Latino Victims	Asian Victims
White82*	.66*	.82*
Black	—	-.12	.07
Latino	-.05	—	.03
Asian	-.07	.00	—

NOTE.— $N = 51$

* Significant at $P < .05$. Entries are Pearson correlations. The .82 correlation in the upper left-hand corner, e.g., depicts the statistical association (across 51 community districts) between the number of antiblack attacks by unknown perpetrators and the number of antiblack attacks by white perpetrators.

attacks (Herek and Berrill 1992; Comstock 1991; Shah and Pease 1992). Clearly, certain victim groups are more willing than others to report their incidents to the police, and for this reason we do not attempt to compare the relative incidence of antiblack, anti-Latino, and anti-Asian crime. On the other hand, because these data were collected by a single administrative unit that takes cases as they are referred by patrol officers, they have the virtue of being subject to similar reporting biases across community districts. It should be pointed out that the results we report below, using police records for all sorts of crime, are very similar to what we find when we restrict attention to incidents involving violent crime, which presumably are more likely to come to the attention of the police.

Unfortunately, we do not have the luxury of being able to compare police statistics on racially motivated crime to those gathered by monitoring organizations analogous to the Anti-Defamation League of B'nai B'rith or the New York City Gay and Lesbian Anti-Violence Project; systematic monitoring efforts simply do not exist for racially motivated incidents in New York City. Similarly, our experiments with newspaper clipping services suggest that even local community papers contain few incident reports and none that escape notice by the police. It would appear that public opinion surveys' assessments of victimization represent the sole means by which to crossvalidate our results. We report the results from two such surveys below. For the present, we note that they corroborate our findings concerning the link between neighborhood composition and racially motivated crime.

MEASURES OF DEMOGRAPHIC CHANGE AND ECONOMIC CONDITIONS

Although we had originally planned to track crime and racial composition on an annual basis, the only reliable demographic data below the borough level come from the decennial census.⁷ Two census indicators are used to assess racial composition and change in each district. A district's prior racial composition is measured by the percentage of non-Hispanic whites in 1980. The measure of racial change for each minority group is the difference between that group's percentage in the 1980 and 1990 censuses. For

⁷ The annual Current Population Survey cannot be disaggregated below the borough level, and the triennial Housing Surveys contain too few respondents per community district to gauge changing racial composition. The lack of information about the timing of demographic change creates some uncertainty about the sequence of cause and effect since some of the racially motivated crimes might have occurred prior to the neighborhood change. Our results would not change materially, however, if we restricted our attention to racially motivated crimes occurring after the decennial census of 1990.

instance, in Brooklyn District 7, Asians comprised 3.7% of the population in 1980 and 10.4% in 1990. Subtracting .037 from .104 gives a .067 change in the Asian population proportion.⁸ Granted, these measures may be operationalized in alternative ways (e.g., ratios of population sizes rather than percentage-point changes; Green 1998) or replaced by other measures of intergroup heterogeneity and mixing. Our experience using such measures is that they reveal similar patterns but are more cumbersome to interpret or to replicate in other settings.

Many potential indicators of white economic vulnerability suggest themselves: poverty rates, median incomes, unemployment rates. As it happens, these indicators tend to be highly intercorrelated so that the choice among them has little bearing on the conclusions one draws. (The correlation between unemployment rates and poverty rates is .95; between unemployment and the log of median income, $-.94$.) After experimenting with various economic indicators, such as median income and poverty rates, we finally settled on unemployment rates for non-Hispanic whites. These district-level data for 1980 and 1990 were obtained from a special tabulation performed by the Bureau of the Census. Using these figures, we constructed measures of both the rate of unemployment in 1990 and change in unemployment over the period 1980–90. The models presented below also explore possible interactions between these indicators of economic conditions and rates of demographic change.

Although community districts are similar in terms of population size, they are not entirely uniform. To show that the results are in no way affected by population differences, we include the natural log of population as a control. This regressor, however, never proves statistically significant, nor does its inclusion alter the effects of other variables.

STATISTICAL MODEL AND DATA ANALYSIS

Allowance must be made for the fact that the observed number of racially motivated crimes is a dependent variable that is restricted to positive integers. Since ordinary least squares (OLS) is inappropriate for analyzing event-count data, we must look to alternative modeling approaches. Poisson regression is widely used to model event-count processes but presup-

⁸ Several studies have used a similar technique to measure demographic change (Marshall and Stahura 1979; Emerson 1994). Other studies have used a measure of percentage black (rather than percentage white) as an independent variable and tested its effects on segregation, voting behavior, and white flight from neighborhoods (Blalock 1957; Giles 1977; McCreary, England, and Farkas 1989; Logan 1981). Admittedly, some of what we are labeling "in-migration" is in fact population change due to racial differences in mortality and fertility.

poses the statistical independence of distinct events, an assumption that would be violated if incidents trigger reprisals or if our model omits causative features of community districts. We therefore model these data using negative binomial regression, which relaxes the assumptions underlying the Poisson model so as to allow for excess variability, or overdispersion, among the event counts (King 1989). The maximum likelihood estimates of the negative binomial parameters are typically more conservative than comparable estimates of Poisson parameters.

The negative binomial model assumes that y_i , the rate at which incidents occur in a community district i during the period 1987–95, has an expected value of λ and a variance of $\lambda\sigma^2$. The expected rate can in turn be modeled as a function of K independent variables ($x_{1i}, x_{2i}, \dots, x_{ki}$). Since λ must be positive, it is conventional to express it in terms of an exponential function:

$$\lambda_i = \exp(\alpha + \beta_1 x_{1i} + \dots + \beta_k x_{ki}).$$

The overdispersion parameter σ^2 represents the factor by which the variance of λ_i exceeds its expectation. Where $\sigma^2 = 1$, the negative binomial model is equivalent to the Poisson. Because σ^2 is assumed to be greater than or equal to 1, we parameterize it as $\sigma^2 = 1 + e^{\gamma}$. Larger values of γ thus indicate more excess variance.

Before embarking on this statistical analysis, let us take a preliminary look at how the incidence of racially motivated crime varies by racial composition across districts. Table 4 reveals a strong monotonic relationship between the white percentage of the population in 1980 and the average incidence per district of racially motivated crimes against Asians, Latinos, and blacks. (Similar conclusions emerge using the white population proportion in 1990 or an average of 1980 and 1990; we use 1980 to facilitate comparison with the models presented below.) This relationship holds regardless of whether one considers all reported crimes or just those by white or unknown perpetrators. The fact that the relationship between these variables is positive runs counter to the predictions of the power-threat hypothesis, which associates white resistance with the numerical superiority of nonwhites. And the fact that the average number of incidents per district always rises monotonically calls into question the tipping point and random-interaction hypotheses, both of which suggest an inverted-U pattern of some kind. Evaluating this bivariate relationship using a negative binomial model leads to the same conclusions: the slope estimates are always positive ($p < .05$), and there is no evidence whatsoever of an inverted-U pattern (i.e., a negative quadratic term). Rather than belabor the point, we henceforth focus our attention on the power-differential and defended neighborhood models.

While the white proportion of the population is an important predictor

TABLE 4
RACIALLY MOTIVATED CRIME IN NEW YORK CITY, 1987-95

	PERCENTAGE WHITE IN 1980							
	All Crimes				Crimes Committed by White or Unknown Perpetrators			
	0-25	25-50	50-75	75-100	0-25	25-50	50-75	75-100
Average no. of antiblack incidents per district	8.2	11.3	22.4	33.1	7.5	10.1	21.3	31.6
Average no. of anti-Latino incidents per district	2.3	3.7	4.6	8.6	1.3	2.6	3.9	8.3
Average no. of anti-Asian incidents per district	1.7	2.4	3.7	5.8	.8	1.6	2.3	5.1
N of community districts	15	9	11	16	15	9	11	16

of antiminority crime, it is not the whole story. Equally crucial is the growth rate of a region's minority population. Rieder (1985) argues in his ethnographic account of mid-to-late-1970s Brooklyn that in-migration of minorities is felt most acutely in homogeneous white neighborhoods, where certain inhabitants, with the tacit approval of others within the community, use various illegal means to thwart residential integration. If this pattern holds generally, we should expect to see not only more racially motivated crime where whites make up the preponderant share of the population but the most acute problems where homogeneous white regions experience a sudden surge in minority population. Conversely, the effects of in-migration will be muted when it occurs in areas that are already integrated. In the limiting case where whites constitute only a small fraction of the population and no longer envision their neighborhood as their racial property or possess the means to enforce such a vision, an in-migration of minorities may actually reduce the number of racially motivated attacks against that group."

In order to test whether long-standing racial predominance in a community influences racially motivated crime, we included the percentage of whites living in each district in 1980 as an independent variable. To gauge the threat posed to the status quo by in-migration of a given racial or ethnic group, we calculated the black, Latino, and Asian population proportion in 1980 and 1990. As noted above, the percentage-point change in each group's population share serves as our measure of in-migration. The critical feature of this model is the interaction between whites' long-standing dominance of a region and the in-migration of Asians, Latinos, or blacks. Created by multiplying the white percentage in 1980 by the in-migration variable, this regressor enables us to evaluate whether the effects of in-migration vary depending on the long-standing proportion of whites in the district. A positive interaction suggests that racially motivated crime becomes most severe when minority in-migration occurs in areas that for several years have had a predominantly white population.

We find all three forms of antiminority crime to be powerfully influenced by the interaction between in-migration and the long-standing predominance of the white population. Consider first the case of anti-Asian incidents (table 5). Looking initially at the main effects, the negative binomial estimates suggest that where the Asian population remained stable between 1980 and 1990, racially motivated crimes were more frequent where whites predominated. (To simplify the illustrations that follow, we shall focus throughout on models that omit population size; the corresponding models that include population produce quite similar results.) Absent Asian in-migration, the base rate for anti-Asian crime in a community district over the period 1987-95 was $e^{(-.06 + .5 * .65)} = 1.3$ in areas where in 1980 whites comprised 50% of the population. Where whites comprised

TABLE 5

EFFECTS OF DEMOGRAPHIC CHANGE ON ANTI-ASIAN HATE CRIME

	Model without Control for Population	Model Controlling for Population
Constant	-.08 (.39)	-6.51 (5.59)
% white population in community district, 198065 (.64)	.66 (.63)
Percentage-point change in Asian population (% Asian 1990 - % Asian 1980)	-9.13 (10.74)	-9.49 (11.07)
Interaction: white population percentage \times change in Asian population	32.98 (15.01)	32.01 (15.34)
Natural log of population55 (.48)
Overdispersion parameter08 (.47)	.10 (.46)
Log likelihood	-93.93	-93.26
Squared correlation between observed and pre- dicted incidence of hate crime55	.54

NOTE.— $N = 51$. Dependent variable = N of anti-Asian incidents, 1987–95. Coefficients are negative binomial regression estimates; SEs are in parentheses.

90% of the population, the corresponding rate is $e^{(-.08 + .9^{*}.65)} = 1.7$. To see the interaction between in-migration and white predominance, note what happens in these two districts when the Asian population increases 10 percentage-points. In the district that had formerly been 50% white, the expected incidence rises to $e^{(-.08 + .5^{*}.65 + 1^{*}-.913 + 5^{*}1^{*}.3298)} = 2.7$. The surge in hate crime is even more dramatic in the 90% white district, as $e^{(-.08 + .9^{*}.65 + 1^{*}-.913 + 9^{*}1^{*}.3298)} = 12.9$. In short, in-migration leads to the sharpest upturn in hate crime when it occurs in predominantly white areas.

This point can be made more concretely with reference to actual community districts. Consider, for example, two districts in which whites comprised approximately 75% of the population in 1980. The Co-op City section of the Bronx experienced no change in Asian population and reported one anti-Asian incident. By contrast, the in-migration of Asians into the Flushing section of Queens was 13 percentage points, the most rapid in New York City. The expected rate of anti-Asian incidents in this district was estimated to be 11.9, and the number of actual reports was 11.

Anti-Latino crimes follow a similar pattern. The main effect for per-

TABLE 6
EFFECTS OF DEMOGRAPHIC CHANGE ON ANTI-LATINO HATE CRIME

	Model without Control for Population	Model Controlling for Population
Constant	17 (.45)	-5.86 (5.07)
White population percentage in community district, 1980	1.61 (.63)	1.48 (.64)
Percentage-point change in Latino population (% Latino 1990-% Latino 1980)	-5.34 (7.43)	-6.29 (7.35)
Interaction: white population percentage × change in Latino population	19.60 (11.29)	22.02 (11.26)
Natural log of population	.	.51 (.43)
Overdispersion parameter	.66 (.37)	.65 (.37)
Log likelihood	-115.00	-114.31
Squared correlation between observed and predicted incidence of hate crime	.54	.55

NOTE: $N = 51$. Dependent variable = N of anti-Latino incidents, 1987-95. Coefficients are negative binomial regression estimates, SEs are in parentheses.

centage white is positive, suggesting that among districts where the Latino population has remained constant, one observes more incidents in traditionally white areas. The main effect for Latino in-migration is negative (though swamped by its standard error), suggesting that Latino in-migration leads to *fewer* incidents in predominantly nonwhite areas. One observes a strong interaction (19.6; $P < .05$) between in-migration and white population percentage, implying that anti-Latino incidents occur most frequently in traditionally white areas where the Latino population has surged (table 6).

Again, the interaction between in-migration and the preexisting white fraction of the population can be illustrated by means of a few examples. The Ridgewood/Middle Village section of Queens reported 22 anti-Latino incidents between 1987 and 1995, more than any other district in New York City. This district had been 92% white in 1980 and saw its Latino population grow by nine percentage points. The Forest Hill section of Queens and the Upper East Side of Manhattan were 84% and 89% white in 1980 but experienced only slight growth in Latino population share. Each reported five anti-Latino incidents. Meanwhile, the Jackson

Heights/East Elmhurst/North Corona district in Queens, which was 43% white in 1980 and witnessed a 12 percentage-point growth in its Latino population share, reported one incident. The only district with a faster growing Latino population share was Washington Heights, which, at 29% white in 1980, reported just one incident.

Because antiblack crime is reported with much greater frequency than anti-Latino or anti-Asian crime, the negative binomial model can be estimated with a higher degree of statistical precision. The results match those obtained earlier, with some nuances. Again, we find a powerful interaction between black in-migration and the size of the preexisting white population ($b = 18.3$, $SE = 4.3$). The difference between antiblack incidents and those directed at Asians or Latinos lies in the size and statistical significance of the main effects. As was true for anti-Latino crime, the main effect for in-migration is negative ($P < .05$), suggesting that blacks become safer as their numbers grow in historically nonwhite areas. The main effect for white population share is particularly strong, suggesting that even among districts where no in-migration occurs, racially motivated crimes are more frequent in traditionally white areas. Assuming the black population remains constant, the expected rate of antiblack crime is 25.6 incidents in a 90% white district, as compared to 13.1 incidents in a 50% white district.

The interaction between in-migration and the size of the preexisting white population is well illustrated by three districts that each experienced a 12 percentage-point growth in the black population between 1980 and 1990. In the Flatbush, Rugby, and Farragut section of Brooklyn, where non-Hispanic whites made up 13% of the population in 1980, only three incidents were reported. In the Wakefield/Olinville section of the Bronx, where whites comprised 34% of the population, 11 antiblack incidents were reported. Meanwhile, in the Flatlands/Canarsie section of Brooklyn, a district where whites comprised 84% of the population in 1980 and whose antipathy toward racial and ethnic outsiders was described in the work *Canarsie: The Jews and Italians of Brooklyn against Liberalism* (Rieder 1985), no fewer than 95 antiblack incidents came to the attention of the NYPD Bias Unit. Because Flatlands/Canarsie accounts for 9.9% of all antiblack incidents, it is a potentially influential observation. It turns out, on further analysis, to have little effect on our results. Replicating the analysis without Canarsie/Flatlands increases the standard errors, as is to be expected due to the withdrawal of information, but the parameter estimates remain more or less unchanged. In essence, the pattern of antiblack crime in the Canarsie section of Brooklyn typifies the pattern characteristic of New York as a whole (see table 7).

The implications of the three sets of negative binomial estimates are summarized in table 8. Tracing the expected rates of bias crime listed in

TABLE 7
EFFECTS OF DEMOGRAPHIC CHANGE ON ANTI-BLACK HATE CRIME

	Model without Control for Population	Model Controlling for Population	Excluding Canarsie
Constant	1.74 (.20)	-1.75 (3.45)	-1.93 (3.83)
White population percentage in community district, 1980	1.67 (.28)	1.64 (.28)	1.65 (.29)
Percentage-point change in black population (% black 1990 - % black 1980)	-4.90 (2.68)	-4.97 (2.62)	-5.18 (3.20)
Interaction: white population percentage × change in black population	18.29 (4.33)	18.15 (4.19)	18.93 (8.11)
Natural log of population	..	.30 (.29)	.31 (.32)
Overdispersion parameter	1.48 (.26)	1.46 (.26)	1.48 (.26)
N	51	51	50
Log likelihood	-178.57	-178.09	-174.04
Squared correlation between observed and predicted incidence of hate crime	.67	.67	.47

NOTE.—Dependent variable = *N* of antiblack incidents, 1987–95. Coefficients are negative binomial regression estimates, SEs are in parentheses

the first column, one sees that incidents occur more frequently as the white population share increases. Tracing the incidence of racially motivated crime across the rows, one gauges the conditional effects of in-migration. Comparing rows makes apparent the fact that the effects of in-migration are more pronounced when minorities move into white strongholds. What matters is not simply the proportion of a given minority group in an area but rather the rate at which this proportion has changed and the demographic circumstances in which this in-migration occurred. Even if a given target group constitutes a growing proportion of the population city-wide, it need not follow that acts of hatred directed against the group will rise over time. The trend hinges on the spatial arrangement of population growth, in particular, the extent to which newcomers cross into areas where whites have traditionally been numerically dominant.

Four further observations warrant mention before turning to the role of economic predictors. First, the same set of demographic predictors have very different effects when the dependent variable changes to an alternative measure of racial interaction: interracial marriage. Using data from

TABLE 8

EXPECTED NUMBER OF RACIALLY MOTIVATED CRIMES IN NEW YORK CITY, 1987-95

GROWTH IN MINORITY POPULATION, 1980-90 (%)	PERCENTAGE WHITE IN 1980				
	5%	25%	45%	65%	85%
Anti-Asian Incidents					
Asian:					
15	0	1	3	9	27
10	0	1	2	5	11
5	1	1	2	3	4
0	1	1	1	1	2
Anti-Latino Incidents					
Latino:					
15	1	2	4	10	25
10	1	2	3	7	14
5	1	2	3	5	8
0	1	2	2	3	5
Antiblack Incidents					
Black:					
15	3	8	20	48	117
10	4	8	17	34	68
5	5	9	14	24	40
0	6	9	12	17	24

a special tabulation of the 1990 census, we find that unlike antiblack bias crime, black-white marriages, for example, are not especially frequent in formerly white enclaves that have experienced a surging black population (Green 1998). On the contrary, the interaction terms that predict higher rates of racially motivated crime against blacks tend to predict lower rates of black-white intermarriage. Intermarriage is well predicted by racial heterogeneity; racially motivated crime is not.

Second, alternative model specifications suggest that in-migration of a particular racial or ethnic group predicts racially motivated crimes directed against only that group. In-migration of one racial group, in other words, does not appear to generate attacks against other minority groups. Replicating our analysis of antiblack incidents, for example, using measures of Latino or Asian in-migration as predictors fails to turn up significant effects. Corresponding patterns hold for anti-Latino or anti-Asian crime.

We note, third, the lack of spatial autocorrelation in our crime data. We have said very little to this point about the fact that our analysis

treats geographically proximal regions as the unit of analysis, and one can imagine complications arising from causal influences flowing across nearby community districts. Augmenting our analysis to include the possibility of spatial autoregression (Doreian 1982) reveals very little evidence of spatial interdependence, and the estimates from the revised models differ from those presented here in only trivial ways (Green 1998).⁹ Again, this pattern of results attests to the geographic specificity of racially motivated crime, at least within the relatively large boundaries of a community district. Controlling for demographic composition and change, an unexpectedly high number of attacks against blacks in one part of Brooklyn, for example, does not seem to spread in epidemic-like fashion to neighboring districts.

The specificity with which racially motivated crime follows ethnic change runs counter to "displacement" theories of intergroup aggression (Dollard et al. 1939; Hovland and Sears 1940). Targets vary depending on the particular circumstances of racial change in each district. In this respect, bias crime in New York City resembles ethnic violence in developing nations. As Horowitz (1985) points out, what from a distance appears to be pell-mell mob violence often on closer inspection turns out to be targeted expressions of intergroup hostility. At the same time, our data are inconsistent with the notion that racially motivated crime merely increases with the supply of potential targets. On average, fewer incidents are directed against a racial minority group when its numbers are large relative to the population. What matters is population growth and the racial context in which it occurs.

Finally, we note that the link we find between neighborhood composition and racial intimidation can be corroborated using survey data. In March of 1987, a telephone survey of New York residents conducted by the *New York Times* asked black and white respondents whether they had been "called by an insulting racial term" or "had an unpleasant confrontation with someone that was based on race" within the past year. Table 9 shows that black respondents residing in what they described as predominantly white areas were much more likely than blacks living in black neighborhoods to report such incidents. (Conversely, white respondents living in what they describe as largely white areas were less likely to report racial discord than whites living in black neighborhoods.) Four years later, a WCBS-TV/*New York Times* survey asked New Yorkers

⁹ Gaussian models of spatial autocorrelation as well as autoregression show no significant spatial patterns (results are not included here, but are available from the first author on request). The latter assumes $Y = \rho Y + \exp(XB) + u$, although similar results are obtained when linear models of the form $Y = \rho Y + XB + u$ are used (see Green 1998).

TABLE 9

SURVEY ASSESSMENTS OF THE RELATIONSHIP BETWEEN RACIAL INTIMIDATION
AND NEIGHBORHOOD COMPOSITION

	RACIAL COMPOSITION OF NEIGHBORHOOD				
	All White	Mostly White	Half and Half/Mixed	Mostly Black	All Black
Survey 1 (% yes):					
Blacks	100 (2)	65 (20)	28 (83)	35 (119)	27 (44)
Whites	25 (127)	23 (310)	29 (131)	41 (17)	67 (3)
Survey 2 (% no):					
Blacks	42 (7)	11 (9)			
Whites	19 (91)	11 (195)			

NOTE.—Survey 1 is the *New York Times* Race Relations Survey conducted in March 1987 (ICPSR 9214), survey 2 is the WCBS-TV News/*New York Times* New York City and Suburban Poll conducted in November 1991 (ICPSR 9861). Tabulations are based on New York City residents. In both surveys, racial composition of the neighborhood was based on the response to the question, "Is the neighborhood you now live in almost all white, mostly white, about half and half, mostly black, or almost all black?" In survey 1, respondents are answering the question, "In the last year or so, has a [black/white] person called you by an insulting racial term to your face?" In survey 2, respondents are answering the question, "Would a black person who doesn't live in your neighborhood feel comfortable walking there after dark or not?"

living in all white or mostly white areas whether "a black person who doesn't live in your neighborhood would feel comfortable walking there after dark." White and black respondents were each more likely to say yes when their neighborhood was all, rather than mostly, white. Granted, the number of observations is small and the measures of crime incidence and risk less than ideal. Nevertheless, the relationships are sufficiently strong and consistent to bolster our confidence in the validity of our ecological analysis.

ECONOMIC CONDITIONS AND RACIALLY MOTIVATED CRIME

Is antiminority crime more common in areas where non-Hispanic whites suffer adverse economic conditions? The answer appears to be no. When the unemployment rate among non-Hispanic whites in 1990 is added to the specifications reported in tables 5–7, unemployment in every case has a weak and statistically insignificant coefficient, while the other coefficients remain intact. Much the same result turns up when the change in

unemployment rate from 1980 to 1990 is substituted as a regressor. Again, the unemployment coefficients are statistically indistinguishable from zero, and the results reported in tables 5–7 remain unchanged.

Experimenting with a variety of different economic indicators (e.g., poverty rates, median income) and specifications (e.g., ratios of white to nonwhite unemployment rates), we searched in vain for evidence linking economic conditions to racially motivated crime. In no case did we find evidence to support the notion that incidents are more frequent in economically depressed areas, nor do we find any statistical evidence of interactions between in-migration and economic vulnerability of whites. This result, too, holds across a wide range of alternative specifications and measures. We find no evidence, for example, to support the view that minority in-migration into areas inhabited by affluent whites (who possess the wherewithal to relocate) is less likely to produce racially motivated crime than in-migration into less affluent white areas. It may be that well-to-do whites are just as likely as other whites to resort to ethnic intimidation in defense of their territory, or it may be incorrect to assume that the affluent can relocate as readily as lower-income people, many of whom may be renters rather than home owners. Either way, our data reveal no simple links between economic circumstances and the quantity of racially motivated crime. (In the interests of conserving space, only a portion of these results are reported in table 10, but the rest may be obtained from our Web site.)

This is not to say that economic evaluations are absent from concerns about racial or ethnic change. Focus group interviews with bias crime perpetrators in Brooklyn suggest that economic grievances and insecurities suffuse the ideology of neighborhood defense (Pinderhughes 1993). On the other hand, we do not find, as have historical studies (Tolnay and Beck 1995; Hepworth and West 1988; see, however, a critical replication in Green et al. 1998), a connection between objective indicators of economic performance and rates of racially motivated crime. It may be that our measures of economic strain are too crude to pick up on the subtleties of economic strain and resource competition, although we do not find stronger links when unemployment rates are replaced by rates of blue-collar employment among whites. If economic vulnerability triggers acts of racial prejudice, the relationship is not apparent from the measures at our disposal.

DISCUSSION AND CONCLUSION

Missing from the extensive literature on the economic and social consequences of residential integration has been systematic analysis of racially

TABLE 10

EFFECTS OF ECONOMIC CONDITIONS AND DEMOGRAPHIC CHANGE ON RACIALLY MOTIVATED CRIME AGAINST BLACKS

	MODEL						
	1	2	3	4	5	6	7
Constant	-1.75 (3.45)	-1.27 (3.51)	-1.27 (3.59)	-1.73 (3.77)	-1.74 (3.49)	-1.68 (3.55)	-1.68 (3.52)
White population percentage, 1980	1.64 (.28)	1.51 (.32)	1.51 (.42)	1.63 (.29)	1.65 (.29)	1.66 (.29)	1.66 (.28)
Percentage-point change in black population	-4.97 (2.62)	-5.11 (2.55)	-4.95 (7.79)	-5.57 (7.95)	-4.97 (2.61)	-4.40 (2.62)	-4.41 (2.62)
Interaction: white percentage \times change in black population ..	18.15 (4.19)	18.53 (4.15)	18.49 (8.60)	18.33 (4.81)	18.11 (4.19)	17.98 (4.18)	18.00 (4.15)
Log of population30 (.29)	.28 (.30)	.28 (.29)	.30 (.32)	.30 (.30)	.29 (.30)	.29 (.30)
Non-Hispanic white unemployment rate, 1990		-2.58 (3.09)	-2.59 (4.08)				
Interaction, non-Hispanic white unemployment \times change in black population			-2.00 (99.25)	7.05 (87.66)			
Percentage-point change in non-Hispanic white unemploy- ment63 (3.43)	.17 (5.29)	
Interaction: change in non-Hispanic white unemployment \times change in black population							
Overdispersion parameter	1.46	1.45	1.45	1.46	1.46	1.45	1.45
Log likelihood	-178.09	-177.71	-177.71	-178.08	-178.07	-177.77	-177.77

NOTE.— $N = 51$. Dependent variable = N of antirblack incidents, 1987–95. Coefficients are negative binomial regression estimates, SEs are in parentheses. SEs reported for model 3 were obtained through bootstrapping because the Hosmer as the maximum likelihood value is not invertible.

motivated crime.¹⁰ Despite the obvious limitations of our data, the results are both nuanced and intuitively plausible. Racially motivated crime appears to coincide with patterns of demographic change, rising where nonwhites move into white strongholds and falling where nonwhites have long resided in significant numbers. This pattern holds in separate analyses of racial incidents directed against Asians, Latinos, and blacks—despite the fact that these groups have established themselves and experienced population growth in different areas of New York City.

As with any ecological analysis, the microlevel mechanisms at work here are open to competing interpretations. One that draws on ethnographic studies of white urban neighborhoods (e.g., Rieder 1985) traces acts of racial intimidation crime to perpetrators' desire to preserve racially homogeneous residential areas and the ways of life that residents associate with them. The admixture of outright racism, nostalgia, and self-interest that contributes to this desire doubtless varies, but the conjunction of this exclusionary sentiment and the tacit support (or active encouragement) of neighbors leads to a heightened propensity for action when racial homogeneity is threatened. More integrated areas experience fewer racially motivated crimes, by this account, because integration (1) gradually undermines the extent to which whites regard a given area as their territory, (2) disrupts the social networks and norms that facilitate attacks against racial minorities, and (3) causes the most virulent opponents of integration to quit the community. Drawing upon a power-differential interpretation, one might argue as well that as white numbers diminish so too does their capacity to attack with impunity.¹¹

Note that this perspective on racially motivated crime says little about reductions in prejudice that occur as a function of integration and, by extension, intergroup contact. Few studies of white attitudes toward minorities find a strong link between racial context and racial attitudes (for a review of this literature, see Sears and Citrin [1997]), and New York

¹⁰ Several studies of interracial crime rates have attributed such crimes to racial hostility (O'Brien 1987). The difference between those crime statistics and the data used in this study is that we possess extrinsic evidence of racial animus (name-calling, racist graffiti, etc.).

¹¹ Our results differ from the expectations of the power-threat hypothesis in ways that may be theoretically informative. Both classic and contemporary studies that found whites to be most reactionary in areas where blacks predominated focused on the Deep South. Unlike New York City, the Deep South bestows on whites a more overwhelming monopoly of power, such that even a locally outnumbered white population may still expect to ride herd over a black majority. Whites living in integrated areas in New York City typically have no such expectation of maintaining or reestablishing white hegemony. Indeed, anticipated loss of control may be what leads the most intolerant whites to take action against minorities who encroach upon white enclaves and to leave when efforts to drive out or subordinate minorities fail.

neighborhoods seem to be no exception.¹² It is doubtful whether the prospect or process of racial integration changes white racial attitudes appreciably (Skogan 1995). The observed link between racial composition and racially motivated crime may be due to the ways in which prejudice is *mobilized* by demographic composition and change. Our results, then, reaffirm the importance of distinguishing hostile racial attitudes from actions (cf. Green and Cowden 1992).

Why the weak association between racially motivated crime and economic conditions? To be sure, our findings run afoul of historical scholarship on racial conflict, but we doubt that the divergent findings stem from the different time periods under investigation. First, the putative connection between cotton prices and lynchings of blacks is methodologically suspect (Green et al. 1998). Second, the violent outbreaks against racial minorities that authors such as Olzak (1989) trace to economic causes involved, to a significant degree, orchestration by labor unions or other organized interest groups. Clearly, economic conditions shape the tactics and vitality of political organizations, whether it be the Nazi Party in Germany during the 1930s or the Ku Klux Klan in the United States during the 1970s (Wade 1987). The question is whether similar dynamics shape the patterns of unorganized mass action carried out sporadically by small groups. Our empirical findings may suggest the importance of distinguishing between coordinated and uncoordinated forms of collective action.

If we are correct to suppose that racially motivated crime emanates not from macroeconomic conditions but rather from threats to turf guarded by a homogeneous group, a number of testable empirical propositions follow. For example, the wave of racist violence that ensued in the aftermath of German reunification and crested in 1992 should coincide not with regional or temporal variations in economic performance but rather with spatiotemporal variations in settlement patterns of immigrants and asylum seekers. We should see, for example, xenophobic activity increase most sharply in the former Eastern states that, in 1991, had very few foreigners but were required by newly adopted policies to accept 20% of the new immigrants; conversely, notwithstanding ongoing unemployment problems and the continuing power differential between majority and minority groups, xenophobic violence should subside after immigration rates

¹² Racial attitudes appear to be only weakly related to neighborhood composition in the 1987 survey of New York City residents discussed earlier. Whites living in all white or integrated neighborhoods hold similar views regarding the amount of discrimination against blacks, the degree to which blacks are to blame for their economic circumstances, and the extent to which blacks are responsible for New York City's crime problems. For contrasting results at a national level of aggregation, see Quillian (1995).

dropped and states were no longer required to accept immigrants in proportion to their state's share of the total German population.

Another testable implication is that as white enclaves within large cities gradually give way to integration, one should see a long-term decline in racially motivated crime within urban areas in the United States. The problem in the short run should grow worse in outlying areas, where homogeneous white neighborhoods remain common. In effect, racially motivated crime should cascade from cities to suburbs, particularly those suburbs in which whites link their social identities to a racially homogeneous conception of their residential communities. At very least, our findings suggest that greater attention should be paid to the causative role of spatial population arrangements and the manner in which they change over time.

It is interesting to speculate that the link between racially motivated crime and turf defense may extend beyond residential turf to racially proprietary conceptions of other spheres of life. What happens, for example, when women for the first time enter male-dominated occupations, military organizations, or universities, or when interracial or same-sex couples challenge social boundaries concerning acceptable intergroup contact? (See Herek and Berrill 1992; Comstock 1991.) The empirical question is not only whether various forms of trespassing precipitate xenophobic responses but whether a similar set of underlying social-psychological mechanisms—salient social identities, vigilant and widespread social networks—produces these behaviors in different settings.

One such setting about which we have said very little is the homogeneous minority neighborhood. Our hesitation stems from the inadequacies of our data. Public information concerning black-on-Asian crime in New York City, for example, turns out to be too sparse to support the kind of statistical analysis done in this essay, though some evidence suggests that anti-Korean attacks are more numerous in cities where this target population is growing (cf. Kim 1996). And what of racially motivated attacks directed against whites? Although there is no direct way to track Jewish populations over time using government statistics, it is telling that anti-Semitic incidents are reported with greatest frequency in areas such as Crown Heights, where the Hasidic population is highly visible and reputed to be on the rise. Incidents in which whites are subject to racially motivated crime (as opposed to incidents ascribed to anti-Semitic or anti-gay motives) seem to be much more complicated. Cross-sectionally, anti-white incidents correlate with the number of antiblack incidents, and temporally these two monthly time series seem to follow a tit-for-tat pattern (Green et al. 1998). This finding raises the interesting possibility that patterns of racially motivated crime committed by dominant groups are, in general, different from minority-on-majority patterns. Questions such as

these suggest a larger research program in which population flows and racial violence are studied in historical and cross-national perspective (e.g., Lieberman 1980; Horowitz 1985) in an effort to develop and test more refined hypotheses about the conditions under which intergroup contact leads to violence.

REFERENCES

- Alba, Richard D., Nancy A. Denton, Shu-Yin J. Leung, and John R. Logan. 1994. "Neighborhood Change under Conditions of Mass Immigration: The New York City Region, 1970-1990." *International Migration Review* 29:625-56.
- Blalock, Hubert M. 1957. "Per Cent Non-White and Discrimination in the South." *American Sociological Review* 22:677-82.
- . 1967. *Toward a Theory of Minority Group Relations*. New York: Wiley.
- Blau, Peter M. 1977. *Inequality and Heterogeneity: A Primitive Theory of Societal Structure*. New York: Free Press.
- Bobo, Lawrence. 1988. "Group Conflict, Prejudice, and the Paradox of Contemporary Racial Attitudes." Pp. 85-114 in *Eliminating Racism*, edited by P. A. Katz and D. A. Taylor. New York: Plenum Press.
- Bobo, Lawrence, and Frederick C. Licari. 1989. "Education and Political Tolerance: Testing the Effects of Cognitive Sophistication and Target Group Affect." *Public Opinion Quarterly* 53:285-308.
- Bobo, Lawrence, Howard Schuman, and Charlotte Steeh. 1986. "Changing Racial Attitudes toward Residential Integration." Pp. 152-69 in *Housing Desegregation and Federal Policy*, edited by John M. Goering. Chapel Hill: University of North Carolina Press.
- Bowling, Benjamin. 1993. "Racial Harassment and the Process of Victimization: Conceptual and Methodological Implications for the Local Crime Survey." *British Journal of Criminology* 33:231-50.
- Carmines, Edward G., and James A. Stimson. 1989. *Issue Evolution: Race and the Transformation of American Politics*. Princeton, N.J.: Princeton University Press.
- Clark, W. A. V. 1991. "Residential Preferences and Neighborhood Racial Segregation: A Test of the Schelling Segregation Model." *Demography* 28:1-19.
- . 1993. "Neighborhood Tipping in Multiethnic/Racial Context." *Journal of Urban Affairs* 15:161-72.
- Comstock, Gary David. 1991. *Violence against Lesbians and Gay Men*. New York: Columbia University Press.
- Denton, Nancy A., and Douglas S. Massey. 1991. "Patterns of Neighborhood Transition in a Multiethnic World. U.S. Metropolitan Areas, 1970-1980." *Demography* 28:41-62.
- De Sena, Judith N. 1990. *Protecting One's Turf: Social Strategies for Maintaining Urban Neighborhoods*. Lanham, Md.: University Press of America.
- Dollard, John, Leonard Doob, Neal Miller, O. H. Mower, and Robert Sears. 1939. *Frustration and Aggression*. New Haven, Conn.: Yale University Press.
- Dorelan, Patrick. 1982. "Maximum-Likelihood Methods for Linear Spatial Effect and Spatial Distribution Terms." *Sociological Methods and Research* 10:243-69.
- Emerson, Michael O. 1994. "Is It Different in Dixie? Percent Black and Residential Segregation in the South and Non-South." *Sociological Quarterly* 35:571-80.
- Farley, Reynolds, Howard Schuman, Susan Bianchi, Diane Colasanto, and S. Hatchett. 1978. "'Chocolate City, Vanilla Suburbs': Will the Trend toward Racially Separate Communities Continue?" *Social Science Research* 7:319-44.

- Galster, George C. 1990. "White Flight from Racially Integrated Neighborhoods in the 1970s: The Cleveland Experience." *Urban Studies* 27:385-99.
- Garofalo, James. 1991. "Racially Motivated Crimes in New York City." Pp. 161-73 in *Race and Criminal Justice*, edited by M. J. Lynch and E. B. Patterson. New York: Harrow & Heston.
- Giles, Micheal. 1977. "Percent Black and Racial Hostility: An Old Assumption Re-examined." *Social Science Quarterly* 58:412-17.
- Giles, Micheal W., and Melanie A. Buckner. 1993. "David Duke and Black Threat: An Old Hypothesis Revisited." *Journal of Politics* 55:702-13.
- Green, Donald P. 1998. "Replication Dataset and Programs: Defended Neighborhoods, Integration, and Racially Motivated Crime." Manuscript. Yale University, Institution for Social and Policy Studies. (Also available at <http://pantheon.yale.edu/~gogreen/ajrsrepl.html>)
- Green, Donald P., and Jonathan A. Cowden. 1992. "Who Protests? Self-Interest and White Opposition to Busing." *Journal of Politics* 54:471-97.
- Green, Donald P., Jack Glaser, and Andrew Rich. 1998. "From Lynching to Gay-Bashing: The Elusive Connection between Economic Conditions and Hate Crime." *Journal of Personality and Social Psychology* 75:82-92.
- Green, Donald P., and Andrew Rich. 1998. "White Supremacist Activity and Cross-burnings in North Carolina." *Journal of Quantitative Criminology*, vol. 14, in press
- Hepworth, Joseph T., and Stephen G. West. 1988. "Lynchings and the Economy: A Time-Series Reanalysis of Hovland and Sears (1940)." *Journal of Personality and Social Psychology* 55:239-47.
- Herek, Gregory M., and Kevin T. Berrill. 1992. *Hate Crimes: Confronting Violence against Lesbians and Gay Men*. Newbury Park, Calif: Sage.
- Horowitz, Donald L. 1985. *Ethnic Groups in Conflict*. Berkeley and Los Angeles: University of California Press.
- Hovland, Carl I., and Robert R. Sears. 1940. "Minor Studies of Aggression: VI. Correlation of Lynchings with Economic Indices." *Journal of Psychology* 9:301-10.
- Kim, Claire Jean. 1996. "Cracks in the 'Gorgeous Mosaic': Black-Korean Conflict and Racial Mobilization in New York City." Ph.D. dissertation. Yale University, Department of Political Science.
- Kinder, Donald R., and Tali Mendelberg. 1995. "Cracks in American Apartheid: The Political Impact of Prejudice among Desegregated Whites." *Journal of Politics* 57: 402-24.
- King, Gary. 1989. "Event Count Models for International Relations: Generalizations and Applications." *International Studies Quarterly* 33:123-98.
- Krueger, Alan B., and Jorn-Steffen Pischke. 1997. "A Statistical Analysis of Crime against Foreigners in Unified Germany." *Journal of Human Resources* 32:182-209.
- Levin, Jack, and Jack McDavitt. 1993. *Hate Crimes: The Rising Tide of Bigotry and Bloodshed*. New York: Plenum Press.
- Levine, Robert A., and Donald T. Campbell. 1972. *Ethnocentrism*. New York: Wiley.
- Liebersohn, Stanley. 1980. *A Piece of the Pie: Blacks and White Immigrants since 1880*. Berkeley and Los Angeles: University of California Press.
- Logan, John R. 1981. "Suburban Racial Segregation as a Nonecological Process." *Social Forces* 60:61-73.
- Marshall, Harvey, and John M. Stahura. 1979. "Black and White Population Growth in American Suburbs: Transition or Parallel Development?" *Social Forces* 58:305-27.
- Massey, Douglas S., and Nancy A. Denton. 1987. "Trends in the Residential Segregation of Blacks, Hispanics, and Asians, 1970-1980." *American Sociological Review* 52: 802-25.
- McCreary, L., Paula England, and George Farkas. 1989. "The Employment of Central

- City Male Youth: Nonlinear Effects of Racial Composition." *Social Forces* 68:55-75.
- Myrdal, Gunnar. 1944. *An American Dilemma: The Negro Problem and American Democracy*. New York: Harper & Row.
- New York Department of City Planning. 1992. *Demographic Profiles: A Portrait of New York City's Community Districts from the 1980 and 1990 Censuses of Population and Housing*.
- New York Times. 1987. *New York Times* Race Relations Survey, March 1987: New York City. MRDF, distributed by Ann Arbor, Mich.: Inter-university Consortium for Political and Social Research.
- O'Brien, Robert M. 1987. "Interracial Nature of Violent Crimes: A Reexamination." *American Journal of Sociology* 92:817-35.
- Olzak, Susan. 1989. "Labor Unrest, Immigration, and Ethnic Conflict in Urban America, 1880-1914." *American Journal of Sociology* 94:1303-33.
- . 1992. *The Dynamics of Ethnic Competition and Conflict*. Stanford, Calif.: Stanford University Press.
- Olzak, Susan, Suzanne Shanahan, and Elizabeth West. 1994. "School Desegregation, Interracial Exposure, and Antibusing Activity in Contemporary Urban America." *American Journal of Sociology* 100:196-241.
- Ottensmamn, John R., and Micheal E. Gleeson. 1992. "The Movement of Whites and Blacks into Racially Mixed Neighborhoods: Chicago, 1960-1980." *Social Science Quarterly* 73:645-63.
- Pinderhughes, Howard. 1993. "The Anatomy of Racially Motivated Violence in New York City: A Case Study of Youth in Southern Brooklyn." *Social Problems* 40:478-92.
- Quillian, Lincoln. 1995. "Prejudice as a Response to Perceived Group Threat: Population Composition and Anti-Immigrant and Racial Prejudice in Europe." *American Sociological Review* 60:586-611.
- Rieder, Jonathan. 1985. *Canarsie: The Jews and Italians of Brooklyn against Liberalism*. Cambridge, Mass.: Harvard University Press.
- Sampson, Robert J. 1984. "Group Size, Heterogeneity, and Intergroup Conflict: A Test of Blau's *Inequality and Heterogeneity*." *Social Forces* 62:618-39.
- Schuman, Howard, Charlotte Steeh, and Lawrence Bobo. 1985. *Racial Attitudes in America: Trends and Interpretations*. Cambridge, Mass.: Harvard University Press.
- Sears, David O., and Jack Citrin. 1997. "Living in the Ethnic Cauldron: Ethnic Context, Social Identity, and Symbolic Politics." Paper prepared for the Social Psychology of Cultural Contact Working Group Conference, Stanford University.
- Shah, Rabindra, and Ken Pease. 1992. "Crime, Race and Reporting to the Police." *Howard Journal of Criminal Justice* 31:192-99.
- Skogan, Wesley. 1995. "Crime and the Racial Fears White Americans." *Annals of American Academy of Political and Social Science* 539:59-71.
- Smith, A. Wade. 1982. "White Attitudes toward School Desegregation, 1954-1980: An Update on Continuing Trends." *Pacific Sociological Review* 25:3-25.
- South, Scott J., and Steven F. Messner. 1986. "Structural Determinants of Intergroup Association: Interracial Marriage and Crime." *American Journal of Sociology* 91:1409-30.
- Suttles, Gerald D. 1972. *The Social Construction of Communities*. Chicago: University of Chicago Press.
- Thernstrom, Stephan, and Abigail Thernstrom. 1997. *America in Black and White: One Nation, Indivisible*. New York: Simon & Schuster.
- Tolnay, Stewart E., and E. M. Beck. 1995. *A Festival of Violence: An Analysis of Southern Lynchings, 1882-1930*. Urbana: University of Illinois Press.
- Tolnay, Stewart E., E. M. Beck, and James L. Massey. 1989. "The Power Threat

- Hypothesis and Black Lynching. 'Wither' the Evidence?" *Social Forces* 67:634–41.
- U.S. Bureau of the Census. 1994. *County and City Data Book*. U.S. Department of Commerce, Washington, D.C.
- Wade, W. G. 1987. *The Fiery Cross: The Ku Klux Klan in America*. New York: Simon and Schuster.
- Willems, H. 1995. "Terror from the Extreme Right." *Terrorism and Political Violence* 7:160–79.

Engineering Growth: Business Group Structure and Firm Performance in China's Transition Economy¹

Lisa A. Keister

University of North Carolina

Business groups have received increasing attention from academics interested in interorganizational relations and their impact on firms. As part of industrial reform, the Chinese government began in the mid-1980s to encourage firms to form business groups with structural characteristics that promised to enhance financial performance and productivity. Using 1988–90 panel data on China's 40 largest business groups and their 535 member firms, the study finds that the presence and predominance of interlocking directorates and finance companies in business groups improved the financial performance and productivity of the groups' member firms. In addition, firms in groups with nonhierarchical organizational structures performed better than firms in hierarchical groups, suggesting that complete integration into a hierarchical organization is not an optimal strategy.

INTRODUCTION

Since 1978, China's government has experimented with market-oriented industrial reform aimed at enhancing the financial performance and efficiency of the nation's enterprises. One of the most dramatic, yet least-studied, components of this effort to engineer industrial growth is the

¹ I am indebted to Howard Aldrich, Heather Haveman, and Victor Nee for helpful comments. I am also grateful to Ronak Brelger, Robert David, Michael Macy, Eugenio Marchese, Rebecca Matthews, Thomas Rawski, Bruce Reynolds, Mary Still, four *AJS* reviewers, and seminar participants at the University of Chicago, Cornell University, University of Michigan, University of North Carolina—Chapel Hill, and Stanford University for helpful comments. Finally, I am grateful to Gary Hamilton and Robert Feenstra for data assistance. This research was supported by grants from the National Science Foundation (SBR-9633121), the U.S. Department of Education, the Cornell East Asia Program, the Cornell Center for International Studies, and the President's Council of Cornell Women. Direct correspondence to Lisa Keister, Department of Sociology, CB#3210, University of North Carolina, Chapel Hill, North Carolina 27599-3210. E-mail: Lisa_Keister@unc.edu

transfer of control of many state-owned firms from government bureaus to newly emerging business groups (*qiye jituan*). Business groups are coalitions of firms, bound together by varying degrees of legal and social connection, that transact in several markets under the control of a dominant, or core, firm (Granovetter 1995). Policy makers studied Japan's keiretsu and Korea's chaebol in preparation for the formation of similar groups in China. In the mid-1980s, the Chinese state began to permit firms to acquire ownership rights in each other and to reduce its own role to that of a shareholder with limited liability and authority (Dong and Hu 1995; Li 1995).² Many of the groups were organized around prior administrative bureaus and most were in manufacturing, though some reorganization and diversification began to occur immediately. By the early 1990s, there were more than 7,000 known business groups in China (Reform 1993). Total 1993 assets of state-owned *qiye jituan* were 1.12 trillion yuan (135.70 billion U.S. dollars), or one-quarter of total state-owned assets (Kan 1996).³ Like the keiretsu and chaebol, China's *qiye jituan* are infused with elaborate interfirm relations, including interlocking directorates, debt relations, and trade ties (Li 1995).⁴

Following the Japanese state's role in the post-World War II formation of the keiretsu, China's state began in the mid-1980s to encourage business groups with certain structural features to emerge.⁵ Chinese officials argued

² In the late 1980s, Chinese securities markets were just emerging, thus the shareholder role was in flux. By the mid-1990s, interactions with foreign firms and the increasing tendency of Chinese firms to list on foreign securities exchanges began to render the meaning of stock ownership in China consistent with its meaning in the West, particularly in the United States (Dong and Hu 1995; Xie 1996). State-owned enterprises related to national security, defense, advanced proprietary technologies, and scarce mineral mining could not be sold to private or foreign investors. A state-owned enterprise in a central industry (energy, transportation, or communications) could be sold, but the state maintained a majority share (Bureau of State Assets Management 1995; Dong and Hu 1995).

³ There are two types of business groups in China: groups of small, often private firms that resemble Taiwan's *guanxi qiye* (Fields 1995), and *qiye jituan*, groups of large, primarily state-owned firms that resemble Japan's keiretsu. I focus on the second type because they are more predominant. Estimates of the proportion of state-owned firms that are members of *qiye jituan* vary with definitions of ownership; 1990 estimates range from 20% to more than 50% (Li 1995).

⁴ The state's intention was to foster economies of scale and to create structures that would ease firms through transition. The chaebol have suffered recently as a result of their size, but Chinese policy makers argue they facilitated development (Li 1995; personal interviews). A difference between Chinese groups and their Asian counterparts is that social relations, while important, played a minor role in group formation in China (though social ties became more important in group reorganization in the 1990s).

⁵ The postwar emergence of the keiretsu was actually a reemergence of the prewar *zaibatsu*, family-centered holding companies. U.S. occupation forces outlawed the *zaibatsu* in 1945, but MITI (the Japanese Ministry of International Trade and Industry)

that when Japan and South Korea were developing, business groups with specific structural traits protected firms from competition, created economies of scale, and enhanced firm performance (Li 1995; PRC 1986). Officials pointed to such features as director interlocks in the *keiretsu* and *chaebol* and group-specific banks in the *keiretsu* as structural components worth emulating (PRC 1980, 1984). Officials have employed "administrative guidance" (including propaganda and asset injections) to increase the likelihood that China's large business groups will develop these same structural features and will, therefore, be less susceptible to the adverse effects that economic shocks and poorly developed markets can have on firms (Nee 1992; PRC 1987).

Existing research proposes that business groups with certain structural characteristics may indeed improve firm performance. This literature provides rich descriptions of business groups in various contexts (Amsden 1989; Fields 1995; Gerlach 1992*b*), but it has largely been limited by data availability to speculation about their performance implications (Aoki 1982; Hamilton and Biggart 1988; Steers, Shin, and Ungson 1989). Recent evidence from Japan demonstrates that *keiretsu* membership reduces variation in firm performance (Lincoln, Gerlach, and Ahmadjian 1996), but even this evidence is insufficient to support claims that there are advantages of specific structural components of the groups. Interorganizational theory suggests that interlocking directorates, a common component of business group structure, will improve performance because they enhance interfirm communication and otherwise reduce transaction costs. However, empirical studies of the interlocks-profits relationship have been inconclusive (Mizuchi 1996; see also Mizuchi and Galaskiewicz [1993] for an excellent review). Research in the United States has generally failed to find a positive effect of interlocks on firm profits, in part because interlocks often form when a firm is in financial decline (Dooley 1969; Richardson 1987). In contrast, research from countries where financial institutions behave differently than they do in the United States finds a positive interlocks-profits relation (Carrington 1981; Meeusen and Cuyvers 1985). As interlocks in Chinese business groups do not result from financial crisis, this case provides a unique opportunity to understand the performance implications of business group structure while clarifying when interlocks matter.

While the bulk of interorganizational relations literature has focused on director interlocks, other interfirm ties may be more consistent pre-

began to resurrect these groups as *keiretsu* as early as 1953 (Gerlach 1992*a*; Johnson 1982). MITI ultimately assembled the groups, supported and guided their core companies, and protected both the groups and their member firms from competition (Miyashita and Russell 1994)

dictors of firm performance (Mizruchi and Galaskiewicz 1993, p. 57). Although it remains to be demonstrated empirically, business groups literature speculates that interfirm credit systems improve performance, particularly where financial markets are weak (Lamoreaux 1994; Granovetter 1995). Finance companies (nonbank, group-specific financial firms) in Chinese business groups are typical of the interfirm finance arrangements discussed in this literature and are thus likely to play an analogous role in improving performance. Transaction cost economics, however, cautions that business group membership benefits firms because the groups economize on control; thus the groups are effective to the extent to which they avoid overorganization by keeping contracts implicit and modes of monitoring informal (Williamson 1985; see also Lincoln et al. [1996, p. 69] for an application of transaction cost ideas to a study of the consequences of business groups). Thus while cooperation via interlocks and financial arrangements may be advantageous, complete integration into a single, hierarchical organization is unlikely to be an optimal strategy (Powell 1990; Powell and Smith-Doerr 1994). The Chinese business groups also provide an opportunity to evaluate these arguments.

My objective is to analyze the effect of business group structure on the financial performance and productivity of the groups' member firms using 1988–90 panel data on China's 40 largest business groups and their 535 member firms. I use 1988–90 data for two reasons. First, the groups did not begin to emerge until the mid-1980s, so these data allow examination of the impact of the business groups in the initial stages of their development. Second, because the groups had become structurally similar by the mid-1990s as a result of the state's promotion of features such as finance companies, these data maximize structural variation. I use multiple indicators of group structure, including indicators of the presence and predominance of interlocking directorates, the presence and predominance of finance companies, and measures of the hierarchical organization of the group. The outcomes I investigate are firm profitability and productivity (measured as output per worker). I evaluate claims about the relationship between the structure of interfirm relations and firm performance and find that coordination through director interlocks and financial arrangements enhances performance but that too much coordination can be detrimental.

BUSINESS GROUP STRUCTURE AND FIRM PERFORMANCE

From the beginning of industrial reform, Chinese firms felt increasing pressure to improve financial performance and productivity, yet the environment in which the firms operated in the late 1980s was not conducive to such improvements. Equity markets were rudimentary: most of the

nation's large domestic banks operated under the aegis of the Central Bank and engaged primarily in government-directed credit extension (Karmel 1994). State funds were limited and were distributed according to social and political, rather than performance, criteria (Li 1995; Spiegel 1994). Private and foreign banks were only permitted to operate under highly constrained conditions, and while Chinese stock markets had begun to develop, trading on these markets remained restricted and provided little capital to firms (Gong 1995). Because product markets were in the initial stages of development, firm access to both inputs and markets for finished goods was limited. Infrastructure limitations and a scarcity of reliable firms specializing in transportation precluded the national distribution of products. In addition, advanced technology was scarce, and increasing competition from foreign firms raised short-term survival concerns and detracted attention and resources from activities that might have led to long-term improvements in financial performance.

Researchers have speculated that business groups with certain structural characteristics may aid firms in overcoming the challenges that accompany development, such as those that faced firms in China in the 1980s (Hamilton 1991; Leff 1978, 1979). Groups that perform banking functions may substitute for more well-developed financial markets and allow firms to obtain otherwise scarce financing. The group may act as a vehicle for mobilizing capital beyond the single family or small group, enlarge the pool from which human resources can be recruited, and allow firms to hire labor where labor markets do not function effectively (Lamoreaux 1986; Leff 1978). Economies of scale may also allow firms to overcome problems associated inefficient product markets, to engage in research and development, and to contend more effectively with foreign competition (Aoki 1982; Granovetter 1995). In addition, the elaborate interfirm relations engendered by the groups may improve the flow of communication among firms, reducing the cost of gathering information and facilitating the diffusion of technological and managerial expertise (Leff 1978, 1979).

Of course, the structure of business groups varies widely among contexts. Japan's *keiretsu* are organized either vertically or horizontally and develop across industries. The *keiretsu* generally include a bank, a holding or trading company, and a diverse group of manufacturing firms (Gerlach 1992a; Lincoln et al. 1992). In contrast, Korea's *chaebol* are typically controlled by a single family or a small number of families and are uniformly vertically organized (Kim 1991). Business groups in Taiwan (*guanxi qiye*) tend to be small, loosely integrated entities characterized by a didactic managerial style as opposed to the authoritarian style common in Korean and Japanese groups (Fields 1995; Hamilton and Kao 1990). Not surprisingly, Chinese business groups have developed their own unique struc-

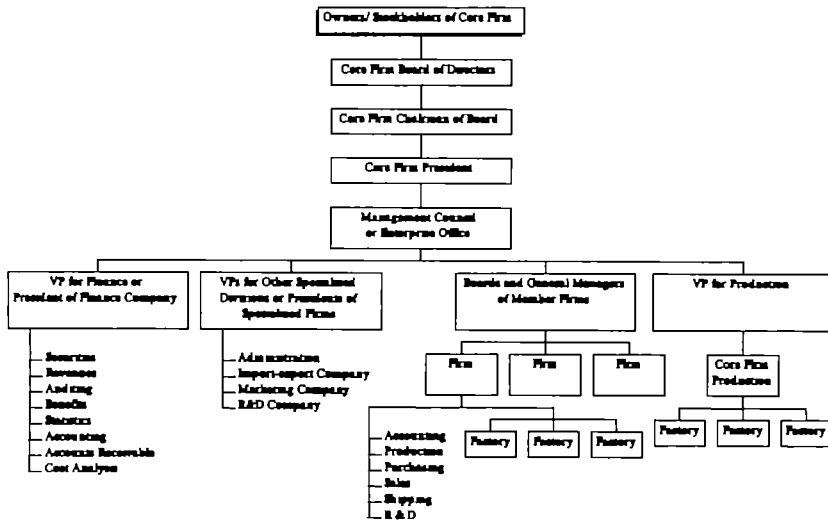


FIG. 1.—The organization of the typical Chinese business group

tures. The groups are large, multi-industry entities with strong ties to the state but not to particular families. Structural variation among the groups has decreased in recent years, in part as a result of the state's efforts to influence group structure. In the late 1980s, however, there was sufficient variation among the Chinese groups in the presence and predominance of interlocking directorates and finance companies, as well as in management structure, to examine their effects on firm outcomes.

Interlocking Directorates

In the 1980s, the board of directors of a Chinese firm was composed of representatives of the firm's owners, including other firms, the business group, and the state. The board oversaw firm management and strategy, chose and oversaw general managers, and made all major financial decisions for the firm (Li 1995; Xie 1996). When a firm entered a business group, partial ownership was transferred to the group's core firm, whose board of directors oversees the activities of member firms. Figure 1 depicts the organizational structure of a typical Chinese business group.⁶ As the

⁶ The figure depicts the typical organizational structure of a Chinese business group, not the typical management structure. Thus, the typical business group portrayed in the figure could be managed either hierarchically (i.e., the core firm could be actively involved in the affairs of member firms) or nonhierarchically.

figure illustrates, the core firm's board of directors is accountable to the shareholders and oversees the activities of the president of the core company.⁷ The president oversees the group's management council or enterprise office, which is composed of the vice presidents and general managers of the group's member firms. The enterprise office is the management office for the core company; it controls production for the core firm and oversees the activities of other member firms (Li 1995; Xie 1996). The finance company, other specialized firms such as an import-export company, the core firm's production division, and the group's other member firms occupy equivalent positions in this hierarchy. The member firm's board of directors contained a representative of the core firm and, in the figure, is located with the member firms below the core firm's management council.

Interlocking directorates are not new to China; they existed in government ministries before reform when a single state representative was assigned to the boards of more than one firm (Schurmann 1965; Xie 1996).⁸ In the business groups, interlocks occur when member firms acquire shares in each other and place representatives on each others' boards. The interlocks have the same functional form as interlocks in other contexts (i.e., an individual occupies a seat on more than one board of directors); however, unlike in the United States, where interlocks are both a source of information (Davis 1992; Haunschild 1993; Mizruchi 1992) and a form of co-optation or monitoring (Aldrich 1979; Dooley 1969; Mizruchi and Stearns 1988), interlocks in China primarily function as an information source for the interlocked firms (Li 1995; SASBG 1995; Xie 1996).⁹ The interlocks allow information about technological advances, market opportunities, innovative strategies, and so on, to pass among firms in the group. Interlocks are not the only source on information available to firms, but they are a central and predominant source. Because interlocks in China are not a form of co-optation, they were viewed differently by Chinese

⁷ Unlike in many Western countries, the Chinese board oversaw the chairman of the board. While there is some separation of ownership and control in business groups, formal ownership was a more direct indicator of control in the early stages of economic transition in China than it is in the West.

⁸ Bendix (1956) argued that the existence of interlocks between ministries and state-owned firms is a practice common to all centrally planned economies.

⁹ My interviews with group managers confirmed this narrower, information-gathering role for interlocks in China and confirmed that managers are aware of most interlocks (see below for data details). Co-optation was unlikely in the early stages of reform because ownership was not well defined and there were few interlocks between groups. In 1988, 40% of China's largest business groups had interlocks; by the mid-1990s, nearly all large groups had interlocks (and finance companies). See table 5 below.

managers during the 1980s than they are generally viewed by Western managers. Chinese managers generally saw them as positive, as they tended to see most forms of social connection among firms in the group (I give more description on interlocking directorates in Chinese business groups elsewhere; see Keister 1998b).¹⁰

Research has shown that interfirm relations affect firm power (Bonacich and Roy 1986), philanthropy (Galaskiewicz and Burt 1991), political behavior (Mizruchi 1992), strategy (Palmer et al. 1987), survival (Miner, Amburgey, and Stearns 1990), and the likelihood of acquisition (Palmer et al. 1995). In addition, researchers argue that interlocks should improve firm performance because they facilitate the flow of information among firms, are a means of co-optation, serve as a monitoring mechanism, and are a reflection of social cohesion (Mizruchi 1996, p. 280). According to resource dependence arguments, informational asymmetries and other uncertainties make corporate environments highly unpredictable (Cook 1977). Interlocks may reduce informational asymmetries by facilitating the flow of information among firms, including collusive information among competitors (Haunschild 1993, 1994; Powell and Brantley 1992), or by facilitating the diffusion of information about innovative practices (Powell 1990; Rogers 1995) and business practices more generally (Davis and Powell 1992; Useem 1984). Another source of uncertainty is resource acquisition: firms depend on resources controlled by other organizations. To minimize dependence and to increase control, firms may use interlocks to leverage resources from other firms, to make others dependent on them, and to monitor the activities of those they control (Pfeffer and Salancik 1978). Interlocks may also serve as an indicator of voluntary relations within which all firms are embedded and may facilitate the unity necessary to carry out joint projects, affect political change, and otherwise manage corporate activity (Granovetter 1985).

Empirical research on the relationship between interlocks and firm performance in the United States, however, has been inconclusive largely because the highly interlocked firms are those in financial decline (Dooley 1969; Mizruchi and Stearns 1988; Meeusen and Cuyvers 1985). In contrast, in countries where the division of labor among financial institutions differs from the United States there may be a positive interlocks-profits relationship (Carrington 1981; Meeusen and Cuyvers 1985). Chinese business groups provide an ideal context in which to examine the interlocks-profits relationship outside the United States. Interlocks in China do not

¹⁰ My interviews confirmed this impression that appears in Chinese literature on interlocks, but my interviews also suggested that, as Chinese firms become more competitive, the meaning of interlocks may converge with their meaning in the West (January 1996, February 1996, March 1997).

develop more readily among firms that are in financial decline; rather they reflect ownership patterns and lines of authority within business groups (Li 1995). These conditions are not unlike those in countries where a positive interlocks-profits relationship has been documented. Moreover, given the high levels of uncertainty in China's transition economy, gaining control over the environment is a central concern for firms (Li 1995). The term *guimo* (scope) refers to the power that group membership gives the firm: it is the collective power (economic, political, and social) of unified action.¹¹ *Guimo* implies economies of scale and access to inputs, financing, markets, and political influence that come with greater size. Interlocks in the business groups increase *guimo* by improving interfirm communication. The interlocks also decrease transaction costs and facilitate the management of resource flows. The role of the interlock is evident in this argument made by the CEO of a Chinese pharmaceutical company: "Interlocks are one of our strongest links to other firms in the business group. Through interlocks, we get ideas about ways to better manage our firm and about technological changes we might otherwise not hear of. . . . [They] also give us an advantage in trade. Other firms know us through the interlocks and are more likely to send the products we need when we need them" (October 1996).¹² In a transition economy, interlocks that are based on political and social connections may facilitate the sharing of rare business information and trading favors among firms. While it is possible that such practices might eventually lead to corruption and irrational decision making, it is likely that during transition, favoritism will aid firms in negotiating ill-developed markets.

Firms involved in the interlocking directorates are not the only companies that will benefit from the improved information flow. Rather, all firms in a business group in which any firms are interlocked will benefit from the presence of the interlocks because member firms are tightly connected through social relations as well as other, more formal relations (e.g., debt relations, personnel exchanges, social ties, political ties). Information passed through the interlocks will continue to spread through the firms' other connections with each other. As a manager in a firm that mines precious metals noticed: "Interlocking directorates are probably the most important ties among firms . . . but managers also have other connections, and news that spreads through the interlocking directorates eventually spreads to all corners of the group because we lend each other money, attend trade shows together, and meet outside of work" (July 1995). Thus, because conditions in China's transition economy resemble conditions in

¹¹ *Guimo* was the primary reason for group membership given by the majority of managers I interviewed.

¹² As the quote indicates, many managers are aware of the benefits of interlocks.

which interlocks have been shown to improve profits, we would expect the following:

HYPOTHESIS 1A. — *Firms in business groups with interlocking directorates will perform better financially and be more productive than firms in business groups without interlocking directorates.*

The logic underlying the argument that interlocks facilitate firm performance through improved information flow also suggests that the advantages of the interlocks will increase as the number of ties increases. The proliferation of the interlocking directorates within a group decreases the time necessary for information to spread to a large number of member firms. The more predominant the interlocks, the greater the number of firms benefiting directly from the information flowing through these ties. Accordingly, in China's transition economy, we should see that firm performance will improve as the proportion of firms in a group with board overlaps increases:

HYPOTHESIS 1B. — *As the proportion of business group member firms tied through interlocking directorates increases, the financial performance and productivity of the firms will improve.*

If interlocks in the business groups improve performance by improving communication among firms, the effect of interlocks should be enhanced by other ties that also improve communication. Joint ventures with foreign firms are an important source of information for firms in a developing economy. Information regarding technological innovations is often spread through joint ventures, and, not surprisingly, firms with joint ventures tend to have performance advantages (Beamish 1993; Chiu and Chung 1993; Schroath, Hu, and Chen 1993). A business group extends the benefits of a single joint venture to all members of the group. If a firm that is not in a business group acquires information from an overseas partner regarding a technological innovation, the information may improve the productivity and performance of the focal firm. If the focal firm is a member of a business group, it is likely that the information acquired through the joint venture will be passed through one or more of the formal or informal connections that exist among member firms. This produces advantages not only for the focal firm but also for the other firms in the focal firm's group. Because information in a business group diffuses through the various interorganizational ties that exist in the group, if a single firm in the group has a joint venture, all firms in the group are likely, over time, to benefit from the information obtained through the joint venture.

Moreover, because director interlocks are a key mode through which information from the joint venture is diffused, the impact of the joint ventures on financial performance will be stronger in groups with director interlocks. An executive in a computer manufacturer observed: "It is dif-

ficult for Chinese-made computers to compete with the faster, smaller computers made by American companies. . . . We need to learn to make smaller chips and to fit more memory in less space. One of the ways we learn to do these things is through joint ventures, our own or those of a fellow business group member. We often get information from [a company in the same group] because our boards are interlocked" (May 1996). Therefore, in China's transition economy:

HYPOTHESIS 1C. — The financial performance of firms in business groups in which any firm has a joint venture will be greater than in groups in which no firms have joint ventures.

HYPOTHESIS 1D. — The effect of joint ventures on firm financial performance will be stronger in groups that have interlocking directorates than in groups without interlocking directorates.

The Finance Company

While the bulk of literature on interfirm relations focuses on interlocks, alternative types interfirm ties may have a more consistent effect on performance (Mizruchi and Galaskiewicz 1993). Researchers speculate that joint ventures, commercial contracts, or financial arrangements may provide greater insight into the functioning of interfirm ties because they vary less across contexts (Amsden 1989; Steers et al. 1989). Economic historians have documented the occurrence of "insider lending" in which a single firm or bank collects and reallocates funds within a group of firms, usually in the early stages of economic development (Gerschenkron 1962; Goto 1982; Lamoreaux 1991, 1994; Munn 1981; Tilly 1966). Similarly, as Japan was developing, business groups that included group-specific banks prospered. These banks began as more informal arrangements, and as the economies developed, aided the group's member firms in financing both short-term projects and activities with more long-term objectives such as research and development (Miyashita and Russell 1994). While this research speculates that a causal relationship exists between the interfirm ties and firm performance, this relationship has not been demonstrated empirically (Mizruchi and Galaskiewicz 1993, p. 57).

Insider lending appears to substitute for a formal financial system and to give firms access to otherwise scarce capital where markets are inadequate at allocating funds (Goto 1982; Lamoreaux 1986). Informal financing arrangements allow funds to be allocated to their highest return uses within a particular group, provides opportunities for diversification, and allows firms to engage in otherwise unaffordable activities. Insider lending can mitigate certain informational asymmetries and reduce transaction costs, allowing firms to gain control over their environments. During de-

velopment, for example, if banks exist, they are likely to be skeptical about unfamiliar potential borrowers. Informal finance arrangements, that are often based on trust among well-acquainted parties, reduce such risks by reducing the amount of information unknown to each party and the costs associated with investigating potential borrowers (Williamson 1981). These arrangements might also provide a vehicle for co-opting resources and thus further reducing environmental uncertainties (Pfeffer and Salancik 1978).

In China in the late 1980s, financial markets were unable to distribute funds efficiently, which left many firms without necessary capital. Firms that were members of some business groups had access to additional financing through the group's finance company (*caiwu gongsi*), a specialized firm that collected and redistributed funds within the group and also obtained funds through state banks on behalf of member firms (Shi 1995). Reformers originally experimented with finance companies in the central industries and later in most other industries (Li 1995). Initially the activities of the finance companies were not monitored, but as their activity expanded regulations were implemented to control lending practices. The finance company enabled the member firms to engage in research and development, to better manage investments both within the group (i.e., investments in other firms that are members of the same group) and outside the group, and, if necessary, to meet short-term operating expenses (for a description of the emergence and functioning of finance companies in Chinese business groups see [Keister 1998a]).

The informational and market-substitute advantages of the group-specific bank suggest that firms in Chinese business groups with finance companies should be advantaged over firms in groups that do not have finance companies, and the more extensive the operations of the finance company, the greater the advantages of this specialized firm.

HYPOTHESIS 2A.—*Firms in business groups with a finance company will perform better financially and be more productive than firms in business groups without a finance company.*

HYPOTHESIS 2B.—*The more extensive the internal financing activities of the finance company, the better the financial performance and productivity of member firms.*

The importance of the finance company to business group member firms is evident in the interaction between finance company activities and joint ventures. Although firms get funding from both joint venture partners and informal financial arrangements, the strong ties among firms in a group make finance company capital more attractive. Thus the effect of joint ventures on performance should be weaker in groups that have access to funds via the finance company:

HYPOTHESIS 2C.—*The effect of joint ventures on firm financial performance will be weaker in groups with finance companies than in groups without finance companies.*

Hierarchical Organization

While interfirm cooperation may be advantageous, it does not follow that complete integration into a single, hierarchical organization is an optimal strategy. Recent research on non-market, non-hierarchical forms of governance acknowledges that such forms are more than hybrids of markets and hierarchies. Stable network forms of organizing, such as business groups, demonstrate that markets are by no means a starting point from which other forms of organizing evolve and that movement toward markets is neither necessary nor desirable (Powell 1990; Powell and Smith-Doerr 1994). Network forms of organizing are effective in facilitating firm performance when monitoring and contractual arrangements are informal (Lincoln et al. 1996; Williamson 1985). Greater integration reduces firm control and restricts the ability of managers to use their knowledge about the needs and abilities of the firm. Therefore, hierarchical organizational forms may increase the firm's transaction costs and negatively affect productivity and performance.

One way to evaluate this claim is to compare performance in hierarchical and nonhierarchical business groups. Some Chinese groups are highly authoritarian: the core firm is actively involved in the day-to-day operations of its subsidiaries. It makes production and personnel decisions in addition to directing more typical matters, such as those regarding corporate strategy. Other groups are more democratic: the core firm allows subsidiaries to manage operations independently. Because the role of the state is reduced to that of a shareholder once the group is established, it can only encourage (rather than require) certain types of management structures to develop in the groups. Thus the management structure in a group develops as a result of state preferences, the preferences of the core firm's managers and board members, and competitive and evolutionary forces. The manager of a steel manufacturer invoked transaction cost ideas in his summary of the potential impact of an authoritarian group management system on the firm: "The business group is very important. You know, *guimo*. We get money and protection from the group, but that doesn't mean a manager in the core firm knows better than I do how to best run my company. The government wants the core firm to be involved in managing member firms, but that is no different from having the state involved. By now bureaucrats ought to realize that was a bad strategy" (March 1996).

Thus, in China's transition economy, we would expect the following:

HYPOTHESIS 3. — *The financial performance and productivity of member firms will be weaker in hierarchical business groups than in nonhierarchical groups.*

RESEARCH DESIGN

Data

To evaluate the claims made in hypotheses 1–3, I collected 1988–90 panel data on the structure of China's 40 largest business groups and the financial performance of their 535 member firms.¹³ I collected the majority of the data during 1995 and 1996 in interviews with core firm managers.¹⁴ While the 1988–90 period saw some retrenchment in economic reform, in part because of events at Tiananmen Square, the formation of business groups continued through this period (curtailment of economic reform at the time of the student democratic movement is discussed by my interviewees; for addition support, see Li [1995]). The core firms are located in 15 provinces and in Beijing, Tianjin, and Shanghai (municipalities that are directly under the jurisdiction of the central government). The member firms are located in all provinces, autonomous regions, and independent municipalities. The business groups in the data set accounted for 68% of the total assets of state-owned business groups in 1990. The member firms are in a variety of industries including manufacturing and ser-

¹³ Because there are missing values on the dependent variables, I use 462 firms in the analyses. The sample includes only large groups (officially, groups with total assets greater than 100 million yuan), thus the results are generalizable only to large groups. I believe that a sample containing only group members is ideal for evaluating my hypotheses, which address relations between group structure and member firm performance. If the hypotheses addressed member vs. nonmember performance differences, it would be necessary to include nonmembers.

¹⁴ I conducted all interviews in Chinese without a translator. I began with 1990 data that Robert Feenstra and Gary Hamilton obtained from the Chinese Economic and Trade Commission; I collected an additional year of data (1988), corrected errors in the original data, and considerably expanded the original number of variables. To maximize accuracy, I personally copied data from the firms' financial statements, spoke with managers formally and informally (out of the plant), and validated the data against other published sources. While errors may still exist, the data appear consistent with other estimates of firm performance (Jefferson and Xu 1991; Naughton 1995). I conducted qualitative interviews in the 40 largest groups, in a random sample of small, medium, and large groups in Shanghai, and in additional groups in underrepresented cities and industries. It would be ideal to compare firm performance in the late 1980s to prereform performance or performance prior to business group membership, but such data are not available.

vices; most firms are former state-owned enterprises, although joint ventures and collective enterprises are also included.¹⁵

Equation Specification and Estimation

To estimate the effects of group structure on firm profits and productivity, I use random effects feasible generalized least squares (GLS) regression equations. The random effects equations decompose the error term to adjust for autocorrelation arising from common firm membership in the same group and for intertemporal correlation of error terms. Because many of the group-level indicators are present in the same groups (e.g., groups with a finance company often have interlocks as well), the test variables are highly correlated. Therefore, I include separate equations (tables 1–3) for each set of test variables (these are not nested models); I also include two equations (table 4) that combine all test variables to demonstrate their joint effect.¹⁶ The equations are of the form

$$Y_{1990} = \alpha + \beta'x_i + \gamma' Y_{1988} + \lambda'G_i + \epsilon_{it}, \quad (1)$$

where Y_{1990} is 1990 profits or output per worker, α is the intercept, x_i is a vector of group- and firm-level control variables, Y_{1988} is a lagged dependent variable, G_i is a vector of group structure variables that test the hypotheses, and ϵ_{it} is the stochastic error term. For the actual estimation of the output per worker equations, all variables (both independent and dependent) are multiplied by the number of workers in 1990 (i.e., the dependent variable is 1990 output, not a ratio).¹⁷

To test hypotheses that propose different processes for firms in groups with different structures (e.g., hypothesis 1c), I model firm performance separately (but in the same equation) for different groups (Greene 1993, p. 582). In these equations, there are two intercepts: one for firms with

¹⁵ A collective is jointly owned by a "guardian" organization (another firm, a social organization, or a government agency) and a rural township or urban municipality. Collectives existed prior to 1978 but were often ignored by the state planning system. Since reform, they have thrived because of their flexible management systems, low labor costs, and ability to retain profits (Oi 1990; Walder 1995).

¹⁶ The random effects model decomposes the error term as follows: $\epsilon_{it} = \alpha_i + \rho_j + \gamma_t + \lambda_{it}$ where ϵ_{it} is the total stochastic component for firm i in time t , α_i is the error component associated with firm i , ρ_j is the component associated with group j , γ_t is the component associated with time period t , and λ_{it} is the stochastic component (Greene 1993). While intertemporal correlation is minimal in panel data, preliminary tests indicated that errors were correlated between 1988 and 1990. Multilevel or hierarchical models use a similar algorithm and return equivalent estimates (Bryk and Raudenbush 1992).

¹⁷ Using the output/worker ratio as the dependent variable did not alter the productivity results substantively.

TABLE 1
1990 PROFITS AND 1990 OUTPUT PER WORKER REGRESSED ON
INTERLOCKING DIRECTORATES

	1990 PROFITS			1990 OUTPUT PER WORKER	
	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept	-1.273** (2.30)	-1.386** (2.52)	..	.630*** (16.43)	.706*** (16.02)
No interlocks	-.278 (.48)
With interlocks	-.203 (.41)
Lagged (1988) profits or output per worker014 (1.59)	.014 (1.60)	.013 (1.55)	.013*** (6.15)	.012*** (6.10)
Measures of group structure:					
Had interlocking di- rectorates (1988) ..	1.057** (2.61)	.	.	.101*** (4.16)	..
% of firms with in- terlocks (1988)	1.960*** (3.22)033*** (3.88)
Had joint ventures (1988)	1.147** (2.55)	.945* (2.10)
With interlocks	2.449*** (3.45)
No interlocks303 (.59)	...	
Group control vari- ables:					
No. of second-tier subsidiaries (1990)	3.375*** (4.04)	3.193*** (3.81)	3.680*** (4.46)	.018** (3.26)	.005 (.82)
No. of third-tier sub- sidiaries (1990)780 (1.26)	.917 (1.49)	.384 (.61)
Firm control variables:					
(log) total assets (1990)286 (.87)	.291 (.89)	.107 (.32)	-.007*** (3.31)	-.012*** (5.10)
Thousands of work- ers (1990)006 (.65)	.005 (.59)	.007 (.87)	.	..

TABLE 1 (Continued)

	1990 PROFITS			1990 OUTPUT PER WORKER	
	Model 1	Model 2	Model 3	Model 4	Model 5
Core firm	3.770*** (3.79)	3.681*** (3.73)	...	-.003 (.50)	.018** (2.65)
With interlocks	8.086*** (5.84)
No interlocks	1.643 (1.57)	.	..
Total sales in group (1990)311*** (3.50)	.304*** (3.44)	.285*** (3.28)
Foreign located ..	-1.226 (1.86)	-1.305* (1.98)	-1.485* (2.30)
Light industry ..	-.801 (.70)	-.877 (.78)	-1.000 (.90)	.087* (2.27)	.062 (1.62)
Adjusted R ²228	.234	.273	.238	.261

NOTE.—Monetary values are in 100 million 1990 yuan (\$12.5 million). Entries are GLS estimates of metric regression coefficients, absolute *t*-statistics are in parentheses. Included in the regression (but not displayed) are dummy variables for having a technology center (a state-supported research division), being in a protected industry, % of profits remitted to the state, location in same province as core firm, and being established since 1978. Data are from 40 Chinese business groups, 462 firms.

* $P < .05$

** $P < .01$

*** $P < .001$

the trait (e.g., interlocks) and one for firms without it. These equations take the form

$$Y_{1990} = (\alpha_1 + \gamma_1'G_{11}) + (\alpha_2 + \gamma_2'G_{12}) + \beta'x_i + \gamma'Y_{1990} + \epsilon_{1i}, \quad (2)$$

where each term is equivalent to the standard equation (1), but where the subscript "1" denotes that the term is for firms in groups with the structural feature of interest (e.g., director interlocks) and the subscript "2" denotes that the term is for firms in groups without the feature.¹⁸

I use two firm-level dependent variables: 1990 firm profits and productivity (output per worker) and standard performance indicators (Meyer 1994; Stickney 1990). Profits are actual profits (revenues less expenses), and productivity is output per worker (output is the actual dependent

¹⁸ The intercept is multiplied by the dummy variable indicating the presence of the structural feature. The use of structural equation explains the absence of the single form of variables that appear to be interactions.

TABLE 2

1990 PROFITS AND 1990 OUTPUT PER WORKER REGRESSED ON FINANCE COMPANY

	1990 PROFITS			1990 OUTPUT PER WORKER	
	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept	-1.807** (2.73)	-.717 (1.41)	..	.556*** (13.36)	.564*** (13.64)
No finance company	-.028 (.80)
With a finance company376 (.67)	..	.
Lagged (1988) profits or output per worker015* (1.71)	.015* (1.80)	.011 (1.32)	.007*** (3.84)	.007*** (3.91)
Measures of group structure					
Had a finance company (1988)	1.423** (2.56)	.		.013* (2.32)	.
Proportion of firms with debt to finance company301*** (4.12)004** (2.44)
Had joint ventures (1988)411 (.79)	1.001* (2.26)
With finance company092* (1.62)
No finance company758* (1.64)
Group control variables.					
No. of second-tier subsidiaries (1990)	3.694*** (4.46)	3.91*** (4.76)	4.745*** (5.89)	.007 (1.40)	.012* (2.41)
No. of third-tier subsidiaries (1990)	1.963** (2.58)	.937 (1.53)	1.641* (2.24)	.036*** (5.20)	.033*** (4.62)
Firm control variables:					
(log) total assets (1990)305 (.93)	.081 (.25)	-.585* (1.70)	-.007** (2.76)	-.004* (2.25)

TABLE 2 (Continued)

	1990 PROFITS			1990 OUTPUT PER WORKER	
	Model 1	Model 2	Model 3	Model 4	Model 5
Thousands of workers (1990)005 (.53)	.004 (.58)	.011 (1.33)	.	..
Core firm	3 896*** (3.90)	3 511*** (3.58)	..	.006 (.67)	-.002 (.24)
With finance company		11.30*** (7.64)	.	..
No finance company	1 750* (1.74)
Total sales in the group (1990)322*** (3.63)	.341*** (3.91)	.305*** (3.60)	-.002* (1.89)	-.001 (1.28)
Foreign located	-1.325** (1.99)	-1.068 (1.65)	-1.291* (2.02)	.068** (2.76)	.076*** (3.12)
Light industry ..	-.986 (.87)	-1.162 (1.04)	-1.325 (1.22)	.062* (1.80)	.077* (2.38)
Adjusted R ²227	.252	.298	.372	.373

NOTE.—Monetary values are in 100 million 1990 yuan (\$12.5 million). Entries are GLS estimates of metric regression coefficients; absolute *t*-statistics are in parentheses. Included in the regression (but not displayed) are dummy variables for having a technology center (a state-supported research division), being in a protected industry, % of profits remitted to the state, and location in same province as core firm. Data are from 40 Chinese business groups, 462 firms.

* $P < .05$

** $P < .01$.

*** $P < .001$

variable, but it is multiplied by the number of workers).¹⁹ Modeling assets turnover (sales/net assets) as the dependent variable produced virtually identical results. I include two interlocking directorates indicators: a dummy variable indicating the presence of interlocks in the business group and a continuous variable indicating the percentage of firms in the business groups involved in the interlocks (both derived from lists of

¹⁹ These are actual, not remitted, profits. I do not use logged profits because there are negative observations. Because I control for firm size and sales, profits can be interpreted as profits per size or per sales. The Shapiro-Wilk statistic (the ratio of the best estimator of the variance—based on the square of a linear combination of the order statistics—to the corrected sum of squares of the variance) indicated that profits and output are normally distributed. The data set does not cover enough time periods to conduct survival analyses.

TABLE 3
1990 PROFITS AND 1990 OUTPUT PER WORKER REGRESSED ON GROUP
MANAGEMENT STRUCTURE

	1990 PROFITS		1990 OUTPUT PER WORKER	
	Model 1	Model 2	Model 3	Model 4
Intercept	-.101 (.18)	-.165 (.30)	1.185*** (16.90)	1.142*** (17.03)
Lagged (1988) profits or output/ worker	.014 (1.57)	.011 (1.33)	.011*** (5.96)	.011*** (5.89)
Measures of group structure:				
Core influenced (in 1988):				
Production decisions of firms	-1.045** (2.54)	..	-.718*** (8.39)	..
Day-to-day operations of firms	..	-1.066** (2.67)	...	-.692*** (8.21)
Group control variables				
Had joint ventures (1988)	1.112** (2.47)	1.192** (2.64)
No. of second-tier subsidiaries (1990)	3.348*** (4.00)	3.332*** (3.98)	.007 (1.24)	.007 (1.23)
No. of third-tier subsidiaries (1990)	.670 (1.08)	.848 (1.37)
Firm control variables:				
(log) total assets (1990)	.307 (.94)	.297 (.91)	-.009*** (4.12)	-.009*** (4.09)
Thousands of workers (1990)	.005 (.54)	.005 (.63)
Core firm	3.746*** (3.77)	3.830*** (3.85)	.008* (1.75)	.009* (1.81)
Total sales in the group (1990)	.313*** (3.51)	.314*** (3.54)
Foreign located	-1.250* (1.89)	-1.245* (1.89)
Light industry	-.876 (.77)	-.875 (.77)	.072* (2.03)	.074* (2.06)
Adjusted R ²	.227	.228	.341	.337

NOTE.—Monetary values are in 100 million 1990 yuan (\$12.5 million). Entries are GLS estimates of metric regression coefficients, absolute *t*-statistics are in parentheses. Included in the regression (but not displayed) are dummy variables for having a technology center (a state-supported research division), being in a protected industry, % of profits remitted to the state, location in same province as core firm, and being established since 1978. Data are from 40 Chinese business groups, 462 firms.

* $P < .05$.

** $P < .01$.

*** $P < .001$.

TABLE 4

1990 PROFITS AND 1990 OUTPUT PER WORKER REGRESSED ON THREE TEST VARIABLES

	1990 PROFITS		1990 OUTPUT PER WORKER	
	Model 1	Model 2	Model 3	Model 4
Intercept	-.743 (.851)	-.101 (.151)	997*** (11.11)	999*** (10.62)
Lagged (1988) profits or output/worker014* (1.92)	.016* (2.26)	.008*** (4.61)	.010*** (3.67)
Measures of group structure				
Had interlocking directorates (1988) .	.718*** 3.12	. (.052*** 3.48)	.976*** (4.13)	... (.108*** 4.21)
% of firms with interlocks (1988)052*** (3.48)	..	.108*** (4.21)
Had a finance company (1988) .	.984*** 4.40	. (.261*** 2.72)	.014*** (3.32)	.. (.240*** 2.43)
% of firms with debt to finance company (1988)261*** (2.72)240*** (2.43)
Core influenced (in 1988):				
Production decision of firms ...	-1.02*** (4.22)	-1.10** (2.16)	-.276*** (3.21)	-.285** (2.32)
Day-to-day operations of firms .	-.992*** (3.31)	-1.14*** (3.11)	-.309** (2.60)	-.295** (2.49)
Had joint ventures	2.31*** (3.36)	1.09** (2.64)	1.19* (1.06)	1.00* (1.20)
Group control variables:				
No of second-tier subsidiaries (1990) ...	1.65*** (3.73)	3.65 (.004)	3.42*** (2.54)	1.05 (.001)
No of third-tier subsidiaries (1990)005 (.312)	2.21 (.000)	.844 (.225)	.505 (.170)
Firm control variables:				
(log) total assets (1990)	2.60*** (2.88)	.628*** (2.31)	.010*** (3.14)	.008*** (3.16)
Thousands of workers (1990) ...	1.23 (.005)	.857 (.000)	.001 (.002)	.002 (.000)
Core firm	2.59** (2.88)	2.34*** (2.62)	.009* (1.06)	.002 (.260)
Total sales in the group (1990)	.332*** (3.90)	.342*** (2.54)	.000 (.019)	.000 (.290)
Foreign located	-2.31*** (3.36)	-2.17** (2.52)	.035* (1.45)	.041* (1.69)
Light industry	-.765 (.720)	-.954 (.920)	.065*** (2.00)	.073*** (2.24)
Adjusted R ²226	.244	.383	.382

NOTE.—Monetary values are in 100 million 1990 yuan (\$12.5 million). Entries are GLS estimates of metric regression coefficients; absolute *t*-statistics are in parentheses. Included in the regression (but not displayed) are dummy variables for having a technology center (a state-supported research division), being in a protected industry, % of profits remitted to the state, location in same province as core firm, and being established since 1978.

* $P < .05$

** $P < .01$

*** $P < .001$

board members for each firm in 1988). I also include two finance company indicators: a dummy variable indicating the presence of a finance company and a continuous indicator of the percentage of firms with debt to the finance company (both are 1998 measures and refer to finance companies that have registered with the government). Two dummy variables indicate whether the group's management structure is hierarchical: one indicates whether the core firm is involved in the production decisions of the member firms (e.g., locating productive inputs, determining output and inventory levels); the other indicates whether the core firm is involved in the day-to-day operations (e.g., personnel matters) of member firms.²⁰ A dummy variable indicates that at least one group member had (foreign) joint ventures in 1988.²¹

A lagged dependent variable in all equations allows interpretation of the coefficients in terms of change in the outcome variable. Because the member firms are spread across a variety of industries, I control for (logged) total firm assets and (logged) number of workers. I indicate whether the firm is in light industry (vs. heavy) to account for capital intensity and remaining industry-specific differences and whether the firm is the core firm.²² I indicate firm integration in the group with a measure of the fraction of the firm's total sales (in 1990) that are in the group. Well-connected firms might perform better not because the group has interlocks, for example, but because the firm has better access than other firms to productive inputs. Because new firms are likely to be more effi-

²⁰ I developed the hierarchy indicators from questions that I asked at least two of the group's managers from different divisions about core firm relations with member firms. These indicators of core firm management strategy across all member firms eliminate particular strategies that might develop in response to the performance of a particular firm. I explored creating a continuous indicator, but the groups fall into two discrete categories (hierarchical and nonhierarchical) on both indicators. I use both measures in the analyses because some firms were hierarchical on one and not on the other. Preliminary tests with the continuous indicator suggest that the relationship between authoritarianism and performance is direct, not curvilinear. However, a more precise measure of authoritarianism is necessary to draw more decisive conclusions about this relationship.

²¹ The presence of joint ventures is a group-level construct because the hypotheses suggest that all members of a group will benefit from the joint ventures of a single-member firm. This variable is not included in the productivity equations for lack of conceptual justification, this was borne out in preliminary tests indicating that this variable did not improve the fit of the productivity equations.

²² I arrived at this industry distinction after extensive examination of the effect of industry on the dependent variables. Other indicators of industry (including both Western and Chinese definitions) and former administrative bureau explained less variance in the outcome variables, and I found no correlation between industry or former bureau and the test variables. Eliminating the core firm from the analyses did not change the results substantively.

cient, I control for whether the firm was established in or after 1978. Geographic controls include an indicator of firm location in a foreign country (which signals access to foreign financing, technology, and management methods) and an indicator of location in the same province as the group's core firm (because proximity reduces the cost of requesting and receiving assistance).²³

I operationalize state involvement in firm affairs with indicators of (1) the presence of a technology center, (2) firm activity in a central industry, and (3) the proportion of profits remitted to the state.²⁴ A technology center is a subsidized research organization; firms with technology centers (generally those dubbed "high tech") have a portion of their expenses for technological research subsidized and receive tax breaks of 30%–50%. The power, steel, iron, automotive, communications, household appliance, and petrochemical industries are central industries; firms in these industries receive state assistance more readily. I use group-level indicators of the number of second- and third-tier subsidiaries to control for size and vertical integration. Second-tier subsidiaries are firms in which a member firm (but not the core firm) has ownership rights; third-tier subsidiaries are firms in which a second tier subsidiaries (but not the core firm) has an ownership interest.²⁵

Table 5 presents descriptive statistics for variables included in the analyses. In 1988, 40% of the groups had interlocking directorates, 40% had a finance company, and 20% of the groups had at least one firm with joint ventures (there is some overlap in having these traits). The indicators of group hierarchy indicate that more than half of the groups are involved in both the production decisions and day-to-day operations of their member firms. Particularly noteworthy are the correlations among the measures of group structure. Consistent with the state's efforts to encourage

²³ Approximately 7% of the firms are located in the same province as the core firm (this low number is not surprising because there is considerable overlap between group membership and membership in administrative bureaus that existed prior to reform and that were not limited by geography). Preliminary tests included additional measures of geographic distance such as an indicator of the average geographic distance between every pair of firms in the group. I also examined urban/rural differences and differences in urban location (e.g., coastal vs. interior). The additional constructs did not improve equation fit.

²⁴ Most firms in the sample did not remit profits in 1988 or 1990.

²⁵ The second- and third-tier subsidiaries draw attention to the issue of group boundary. I include only the 535 firms in which the core firm has an ownership interest for two reasons. First, second- and third-tier subsidiaries are not considered members of the business group by the definition given above. Second, hypothesis 3 concerns the influence of the core firm's management style on member firms, and the core firm has no measurable influence over firms owned by its subsidiaries.

MEANS, SDs, AND ZERO-ORDER CORRELATIONS

	Mean	SD	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1990 profit	20	8.77	.249*	.234*	.015	.101*	.091*	-.129*	-.127*	.114*	.100*	-.005*	.160*	.003	.215*	.092*
1990 output	4.37	13.64	.079	.027*	-.064	.137*	.077	-.214*	-.186*	.128*	.298*	-.114*	.171*	-.009	.435*	.361
1 1988 profit	16	6.67		.687*	.159*	.047	.063	-.015	-.070	.097*	.010	.128*	.134*	-.035	.082	.064
2. 1988 output	3.74	10.92			.158	.081	.098	-.073	-.112	.126	.108*	.114	.101	.012	.161	.107
Measures of group structure																
3. Joint ventures (1 = yes)	.200	.405				-.134*	.498*	.011	.056	.042	-.065	.349*	-.001	.555*	-.099*	-.165*
4. Interlocks	.400	.496					.577*	-.782*	-.625*	.876*	.088*	-.038*	.252*	.041	.109*	.096*
5. Finance company	.400	.496						-.505*	-.421*	.391*	.050	.421*	.314*	.119*	.092*	-.042
6. Core affects production	.652	.483							.798*	.796*	-.096*	.251*	-.295*	-.129	-.136*	-.060
7. Core affects day-to-day	.643	.494								-.692*	-.104*	.171*	-.295*	-.130	-.129*	-.048
8. % of firms in Interlocks	.195	.295									.094*	-.037*	.278*	-.037	.092*	.039
9. % of firms with debt to finance company	.24	.248										-.004	.016	-.042	.510*	.345*
Group control variables																
10. No of direct subsidiaries	32.57	24.83														
11. No. of second-tier subsidiaries	54.15	47.30											-.041*	-.089*	.128	.114*
12. No of third-tier subsidiaries	40.25	48.98												-.038*	.134	.101*
Firm control variables																
13. Total assets	3.07	14.73													-.035	.012*
14. Thousands of workers	9.48	31.96														.161*

NOTE.—Monetary values are in 100 million 1990 yuan (\$12.5 million). Profits, output, total assets, and number of workers are measured at the firm level, $N = 535$. All other variables are measured at the business group level, $N = 40$.

* 1988 values, otherwise 1990 values

* $P < .05$.

the formation of business groups with certain structures, many of the test variables are present in the same groups. In keeping with the idea that the presence of interlocking directorates and a finance company are positively correlated to financial performance, the correlations between these variables and firm profits are positive. Likewise, consistent with the argument that an authoritarian management negatively impacts performance and productivity, the management structure indicators (variables 6 and 7) are negatively correlated with profits and output.

RESULTS

Interlocks Improve Performance and Productivity

A remaining controversy in the sociology of organizations is the effect of director interlocks on firm performance. Interorganizational theory suggests a positive interlocks-profits effect, but empirical studies have been inconclusive (Mizruchi 1996). Table 1 presents GLS estimates of the equations including the interlocking directorates test variables. Consistent with hypothesis 1a, my analyses revealed unambiguously that, in Chinese business groups, interlocking directorates have a positive effect on firm performance and productivity. The presence of interlocks in a business group has a positive effect on the profits (model 1) and productivity (model 4) of the group's member firms (model 1). Moreover, it is clear that the more predominant the interlocks within a business group, the greater the profits and productivity of the member firms (models 2 and 5), as hypothesis 1b proposed.²⁶ Taken together, these results demonstrate not only that interlocking directorates matter but also that they matter for all firms in a group in which any firms are linked through director interlocks.²⁷

²⁶ The data provide strong support for the proposed direction of causation. Because the independent variables are measured in 1988 and the dependent variable in 1990, the coefficient estimates provide some support for the proposed causal direction. Moreover, in 1988, there was no statistically significant difference between the profits and productivity of firms in groups with the test variables and those in groups without the test variables. For example, there was no significant difference in profits or productivity for firms in groups with interlocks and those in groups without interlocks. The same was true for groups with and without a finance company and a hierarchical organizational structure. Yet, both the 1990 profits and productivity of firms in groups that had interlocks in 1988 and those in groups that had a finance company in 1988 were significantly greater than the profits and productivity of those in groups that did not have these traits in 1988. In contrast, the 1990 profits and productivity of firms in groups that were hierarchical in 1988 were significantly lower than those that were not hierarchical in 1988.

²⁷ Including an indicator of focal firm membership in interlocks did not improve equation fit. Extensive investigation of the relationship between interlock predominance and performance revealed that the relationship is strictly increasing; i.e., a threshold effect (Bunting 1976) was not apparent. There was also no evidence of an effect of density of interlock ties.

Why is the relationship between interlocking directorates and firm profitability and productivity unambiguous in these data while prior research has produced mixed results? The answer, in part, rests on differences in the context in which this research and prior research have been conducted. Existing research into the impact of interlocks on firm performance has produced mixed results because it has been conducted primarily in the United States, where firms often add bank members to their boards of directors during financial crises (Dooley 1969; Richardson 1987). My regression results are strong because interlocks in Chinese business groups do not form primarily when firms are in financial crisis. When the Chinese state is influential in organizing a group, officials deliberately combine profit- and loss-making firms into the same groups (Li 1995). The rationale is that the profitable firms will absorb the losses of the other firms; one implication of this practice is that there is not a performance-membership relation in these groups. My interviews confirm this: when groups participate in selecting their members, 82% of firms are chosen for strictly functional reasons (i.e., the firm has some capacity the group needs). The next 11% are chosen because their managers have social or political ties to someone in the group.

The mechanism by which interlocking directorates affect performance is best illustrated with an example relayed to me by the member of the board of directors of one of China's major airlines (also one of the country's largest business groups) in February 1996. Shortly after the group formed in the mid-1980s, a change in airline legislation in the United States would have made it possible for one member of the group to expand considerably into the U.S. market. A manager in another firm learned of the change during a trip to the United States and mentioned it as part of his report to his firm's board of directors. A board member, who heard the report and who also held a seat on the board of the firm that might benefit from the information, recognized the importance of the information. He relayed it to the relevant manager in the other firm, whom he knew from meetings of that firm's board. The second firm was able to act on the information and expand its sales. Perhaps the role interlocks play in relaying information is their single greatest advantage in China.

The mechanism by which interlocks affect performance is also clarified by the interaction between interlocks and joint ventures, another source of information in the groups. Prior research predictably finds that firms benefit from joint ventures because contact with foreign firms often transfers technology and management expertise to the Chinese firm (Beamish 1993; Chiu and Chung 1993; Schroath et al. 1993). When a single firm in a business group has a joint venture, all the firms in the group benefit because the firms in the group are tied to each other in various ways.

Information entering the group through the joint venture is passed among member firms through interlocking directorates and other formal and informal linkages. Model 3 of table 1 includes the results of the structural equation model in which the profits of firms in groups with and without interlocking directorates are modeled with separate regression lines. As hypotheses 1c and 1d proposed, the effect of joint ventures is positive and highly significant for firms in groups with interlocks but not significant (and relatively small) for firms in groups with no interlocks.²⁸

An alternative explanation of the positive effect of interlocks on profits and productivity involves the role political and social ties play in the groups. Most interlocks develop on the basis of ownership connections, but in those based on political and social connections, the interlocks may be more than a conduit for technological information. They may instead facilitate nepotism, the sharing of rare business information, and the trading of favors among firms. Additional research on the mechanism by which interlocks in the business groups facilitate performance could clarify this.

Note that in models 1–3 in table 1, the effect of lagged profitability is not significantly different from zero. Preliminary tests demonstrated that while lagged (1988) profits predicted current (1990) profits before the interlocking directorate variables were added, their effect was reduced by the addition of the interlock constructs. This is, in part, a testament to the importance of interlocking directorates in business groups because it indicates that interlocks are more influential even than prior performance. In addition, the weakness of the lagged dependent variable may be the result of dramatic changes underway in Chinese industry: as a result of economic reform, firm behavior (and therefore financial performance) has been changing considerably every year. Under such circumstances, prior performance is a weak predictor of current performance. In contrast, lagged output is a strong predictor of 1990 output (table 1, models 4 and 5). In both of these models, however, the intercept is also highly significant, suggesting that omitted variables still explain a large amount of variation in the dependent variable. Omitted variables in these equations and those discussed below might include measures of firm flexibility, group unity, and manager competence.

Finance Companies Improve Performance and Productivity

Literature on economic development has argued that business groups improve access to financing (and thus firm performance) when they contain

²⁸ Because information passes among firms through various interfirm linkages, similar relationships between joint ventures and the presence of other linkages (e.g., trade

a group-specific banking company. Indeed, the results presented in table 2 demonstrate that both profits and productivity are greater for firms in Chinese business groups with finance companies, as proposed by hypothesis 1a. Likewise, consistent with hypothesis 1b, the more extensive the operations of the finance company, the greater the performance and productivity advantages enjoyed by the firm. Model 1 demonstrates that the presence of a finance company has a positive effect on these firm outcomes. In fact, the increase in profits between 1988 and 1990 for firms in business groups with a finance company was 1.4 million yuan greater than the change in profits for firms in groups without finance companies. Model 2 demonstrates that the proportion of firms with debt to the finance company is also a positive influence on firm performance.

In the late 1980s and early 1990s, Chinese firms sought financing from numerous sources, and foreign joint ventures often provided funds that were not available domestically. Accordingly, the presence of joint ventures has a positive relationship with firm profits (though the coefficient is not significant in model 1). Model 3 explores the relationship between joint ventures, the presence of a finance company, and profitability in more depth. In this structural equation model, the variable indicating that the presence of joint ventures for firms in groups that had finance companies in 1988 is positive and significant, while the same indicator for firms in groups without finance companies is positive and significant, but it is also significantly less. This supports hypothesis 2c, which suggests that firms will seek financing from the group's finance company first and then utilize nondomestic sources that are more difficult to obtain and less predictable in the long run.

As the CEO of the core firm of a medium-sized business group in Shanghai argued when I interviewed him in July 1995, firms in China simply do not have adequate access to financing on their own. The first time I spoke with this manager, he was the president of one of the group's most prosperous member firms, and he was dubious that finance companies would have an impact on firm outcomes (in fact, he suggested that my research was not going to yield much useful information). The second time I spoke with this manager, over a year later, he had become the CEO of the group's core firm. In his new position, he had much more contact with his group's finance company, as well as with the finance companies of other groups. His attitude was completely changed, and he argued that because many firms simply cannot acquire adequate credit independently, the finance companies are indispensable, particularly for firms that may

ties, personnel exchanges) should exist. Modeling this possibility, I indeed found similar results. Space considerations prevent the display of these (redundant) results.

still be struggling to rid themselves of the legacy of inefficiency left behind by central planning.

The ability to improve the efficiency of operations is crucial in the transition from state socialism, a system that bred large, inefficient business enterprises. Survival in postsocialism requires the firm to remake drastic changes to production, management, marketing, and nearly all other aspects of corporate operation. Yet, as has been the case in China since the beginning of reform, it is likely that environmental challenges and economic shocks associated with development will impede the realization of these changes. The argument that the business group finance company eases this transition by centralizing the capitalization process for a group of firms is supported by the empirical results and was also evident in my interviews with managers. The 1988–1990 increase in productivity of firms in business groups with a finance company was significantly greater than that of firms in groups without a finance company, shown in the results of models 4 and 5 in table 4. Likewise, the majority of managers in business groups argued that the finance company substitutes for a more formally developed financial market.

Hierarchy Hinders Performance and Productivity

Innumerable managers remonstrated during my interviews that authoritarian group managers were no different than the bureaucrats who oversaw the operations of these firms prior to industrial reform. My results, shown in table 3, confirm the interviews: the financial performance and productivity of member firms are weaker in hierarchical groups. Model 1 includes the variable indicating that the core firm influenced production decisions, while model 2 includes the variable indicating that the core firm influenced day-to-day operations. Both indicators of the core firm's involvement, proxies for the management style of the core firm, are negative and highly significant lending support for hypothesis 3. Model 3 includes the dummy variable indicating that the core firm is involved in the production decisions of the member firms, and model 2 includes the dummy variable indicating that the core firm is involved in the member firms' day-to-day operations. Again, these results demonstrate that management style is an important predictor of productivity as both test variables are negative and significant. These findings are consistent with transaction cost economics arguments that suggest that interfirm networks positively influence firm outcomes when monitoring and contractual arrangements are informal and that complete integration may not be desirable. That is, while interlocking directorates and interfirm financial ties improve firm performance and productivity, it cannot be assumed that even more integration among firms would necessarily continue to

improved performance. Indeed the results presented in table 3 demonstrate that greater interfirm coordination (measured here as core firm activity in firm operations) can be detrimental to firms.

However, this finding is at odds with conventional understandings of the role business groups played in the development of Japan and Korea. In both the *keiretsu* and in *chaebol*, the core firm managed member firms in a highly authoritarian manner. As Steers et al. (1989, p. 47) point out, "Korean CEOs are seldom challenged, however politely; their decisions are absolute." Decision making in Japanese business groups is much more democratic (and has become more democratic in recent decades) but, relative to the management of groups in countries such as Taiwan, remains relatively authoritarian. The authoritarian model is considered ideal by Chinese officials who are guiding the formation of business groups in their country because it is based on the experiences of Japan and Korea, and because a more authoritarian approach is consistent with Chinese management trends over the past four decades. Yet my results clearly indicate that firms in business groups with less authoritarian management styles perform better financially and are more productive. The manager of a petrochemical firm that is managed by an authoritarian core firm summarized this effect: "Before reform, the government ran enterprises with an iron hand. Now the core firm tries to do the same. Economic reform is supposed to mean change, but this is no change. It is the same system with a new manager. Our firm cannot change until we have the resources to change, but also we need the flexibility to make our own decisions. We cannot do this if the core company must approve every decision we make" (March 1997).

By the mid-1990s, the business groups included in these analyses had become structurally similar. Nearly all (more than 90%) had interlocking directorates and a finance company. By contrast, fewer groups were rigidly hierarchical (in 1995, fewer than 40% of core firms were involved in the production decisions and day-to-day operations of firms, compared to about 65% in 1990). One explanation for this structural convergence is state pressure for the groups to adopt certain structural features, such as interlocks and finance companies. Alternatively, it is possible that the more efficient governance structures were being selected by emerging market mechanisms and competitive pressures. Both arguments have merit, and in the case of interlocks and finance companies provide an instance of a situation in which state and market pressures are operating in the same, optimal direction. In the case of the degree to which the groups are hierarchically arranged, policy makers and markets seem to be pushing in opposite directions, and markets appear to be winning. Reformers, drawing on experience in Japan and Korea, have encouraged strong core firms in the groups. The negative performance impact of

this type of management and the tendency of groups to move away from authoritarian management suggests that market selection of non-hierarchical organizational forms is leading to more democratic governance.

On a technical note, the models presented in tables 1–3 use separate equations to demonstrate that three aspects of business group structure (interlocking directorates, a finance company, and the hierarchical structure of the group) affect member firm financial performance and productivity. Table 4 presents estimates of four models that combine indicators of each structural element into the same equation. The results in this table demonstrate the high joint significance of the test variables and demonstrate that the individual effects of the test variables persist when these variables are used to simultaneously predict the dependent variables. Model 1 regresses 1990 profits on the dummy variable indicators that the group had interlocks and a finance company, that the core firm influenced production decisions, and that the core firm influenced day-to-day operations. Model 2 includes the continuous indicators of the percentage of firms involved in the interlocks and the percentage of firms with debt to the finance company along with the dummy variable indicators of whether the group is hierarchically organized. Models 3 and 4 regress output per worker on the same sets of test variables. In all cases, the coefficient estimates for the test variables are of the same relative magnitude as those in the separate equations presented in tables 1–3, although (perhaps because of correlations among the test variables) standard errors tend to be higher and some coefficients for control variables are considerably larger or smaller than in the separate equations.

CONCLUSION

My objective was to examine the effect of business group structure in China's transition economy on the financial performance and productivity of the groups' member firms. I started by observing that one of the most profound components of China's industrial reform has been the reorganization of firms into business groups, a process that began in the mid-1980s. I argued that using multiple indicators of group structure—including indicators of the presence and predominance of interlocking directorates, informal finance arrangements, and the hierarchical organization of the group—would clarify the role that business groups play in determining firm performance in China and would also inform understanding of the importance of these groups and interorganizational relations more generally. Using data on China's 40 largest business groups and their 535 member firms in 1988–90, I evaluated a series of hypotheses drawn from the literature on interfirm relations and found strong support for each hypothesis.

I found strong support for hypotheses that anticipated a positive relationship between the presence and predominance of interlocking directorates and firm performance. Earlier empirical studies of interorganizational relations have focused on the impact of interlocks on firm profits, primarily because data on both the outcome and explanatory variables have been readily available. However, this research has produced inconclusive results because financially troubled firms, in the contexts in which the research has been conducted, are more likely to be interlocked. The inconclusive results in this literature do not indicate that interorganizational relations do not affect firm outcomes; instead they demonstrate that the available data are inadequate for testing propositions about interlocks and performance. By contrast, Chinese firms that are in financial decline are not more likely than other firms to be involved in interlocks. In China, interlocking directorates improve information flow among firms and thus reduce the cost to an individual firm of obtaining and processing information. The result is improved performance over comparable firms also in business groups that do not have access to the interlocks.

To strengthen this finding, I also examined the effects of informal finance arrangements—another interorganizational relation present in the business groups—on firm performance. As Mizruchi and Galaskiewicz (1993, p. 57) note, because literature on interlocks and performance has been controversial, researchers need to begin considering how alternative interorganizational relations affect firm performance. This is possible in Chinese business groups because they constitute a unique type of interfirm network that is defined not by a single tie among firms (such as interlocking directorates) but is interwoven with innumerable linkages, including financing arrangements, personnel exchanges, production agreements, and social ties. I found that interfirm financing arrangements improve both firm profitability and productivity and that the more extensive these relations, the greater the benefit to the firm.

Evidence that interorganizational cooperation improves performance does not, however, imply that complete integration into a single, hierarchical organization is an optimal strategy. Because an extremely large corporation reduces managerial flexibility, we would expect an intermediate (between markets and hierarchies) organizational structure to be not only stable but also most beneficial to firms (Powell 1990). My results—particularly the finding that firm performance declines as central control in the business group increases—provide a limited amount of support for this notion. Because the analyses presented here do not include firms that are not business group members, it is unclear how nongroup members would differ. Within the group, however, the results suggest that more control is not necessarily better. Over much longer time periods, the possibility for movement toward the extremes always exists, particularly as environ-

mental conditions change. In the case of the *keiretsu*, Gerlach (1992a) has raised similar questions about whether this organizational form might be in decline as the Japanese economy has moved beyond the development stage. My results indicate, however, that under at least some conditions (specifically, moderate to high amounts of uncertainty resulting from poorly developed markets and declining state control), firms subjected to less authoritarian organization perform better.

The formation of business groups has been one of the most profound components of China's efforts to engineer industrial growth. The deliberate disengagement of formerly state-owned enterprises from the command of administrative bureaus is, in part, a result of the perception that business groups with specific structural characteristics protected firms in other countries from the shocks and challenges of development. Research on business groups has argued that groups with certain structural characteristics may protect firms from competition, allow them to take advantage of economies of scale, and substitute for more formal financial markets (Hamilton 1991; Leff 1978, 1979). Comparisons between the Chinese case and the emergence of business groups in Japan and Korea following World War II, are inevitable given the geographic proximity of these countries and the deliberate efforts of the Chinese state to reproduce the groups of its Asian neighbors. Little empirical evidence exists to support claims that elements of group structure provided advantages to firms in Japan or Korea (Lincoln et al. [1996] is an excellent exception); my results support the idea that groups with certain structures can improve firm performance.

Extensions of this research to other contexts would clarify the exact conditions under which the relations demonstrated in this study hold. Future research might also expand the scope of the sample to include small and medium-sized groups. My interviews suggest little difference between these and the large groups, but more extensive research could investigate in greater depth the differences that might exist. Future research might investigate more closely the role that a firm's position in the network (e.g., its centrality or connectedness) plays in determining firm outcomes. On the micro level, the processes that lead to the development of interfirm ties might also provide insight into the degree to which the firms are responding to market pressures versus being led by path dependence. An analysis of the emergence of dyads and triads of interfirm exchange ties would answer such questions. Finally, the successful diffusion of the business group concept to China indicates that firms in Vietnam, Eastern Europe, and others undergoing the transition from socialism might benefit from the deliberate formation of business groups with specific structural characteristics.

REFERENCES

- Aldrich, Howard E. 1979 *Organizations and Environments*. Englewood Cliffs, N.J.: Prentice-Hall.
- Amsden, Alice H. 1989 *Asia's Next Giant: South Korea and Late Industrialization*. New York: Oxford University Press.
- Aoki, Masahiko. 1982. "Business Groups in a Market Economy." *European Economic Review* 19:53-70.
- Beamish, P. W. 1993. "The Characteristics of Joint Ventures in the People's Republic of China." *Journal of International Marketing* 1:29-48.
- Bendix, Reinhard. 1956. *Work and Authority in Industry*. Berkeley: University of California Press.
- Bonacich, Phillip, and William G. Roy. 1986. "Centrality, Dominance and Interorganizational Power in a Network Structure: Interlocking Directorates among American Railroads, 1886-1905." *Journal of Mathematical Sociology* 12:127-35.
- Bryk, Anthony S., and Stephen W. Raudenbush. 1992. *Hierarchical Linear Models: Applications and Data Analysis Methods*. Newbury Park, Calif.: Sage.
- Bunting, David. 1976. "Corporate Interlocking: Part III—Interlocks and Return on Investment." *Directors and Boards* 1:4-11.
- Bureau of State Assets Management. 1995. "Tentative Regulations on Ownership Transfer of State Enterprises." *Shanghai Economic Daily*, January 1, p. 1.
- Carrington, Peter J. 1981. "Horizontal Co-optation through Corporate Interlocks." Ph.D. dissertation. University of Toronto, Department of Sociology.
- Chiu Lee-in and Chin Chung. 1993. "An Assessment of Taiwan's Indirect Investment toward Mainland China." *Asian Economic Journal* 7:41-70.
- Cook, Karen S. 1977. "Exchange and Power in Networks of Interorganizational Relationships." *Sociological Quarterly* 18:62-82.
- Davis, Gerald F. 1992. "Agents without Principles? The Spread of the Poison Pill through the Intercompany Network." *Administrative Science Quarterly* 36:583-613.
- Davis, Gerald F., and Walter W. Powell. 1992. "Organization-Environment Relations." Pp. 315-76 in *Handbook of Industrial and Organizational Psychology*, edited by Marvin D. Dunette and Leaetta M. Hough. Palo Alto, Calif.: Consulting Psychologists Press.
- Dong Jie Lin and Hu Jie. 1995. "Mergers and Acquisitions in China." *Federal Reserve Bank of Atlanta Economic Review* 80:15-29.
- Dooley, Peter C. 1969. "The Interlocking Directorate." *American Economic Review* 59:314-23.
- Fields, Karl J. 1995. *Enterprise and the State in Korea and Taiwan*. Ithaca, N.Y.: Cornell University Press.
- Galaskiewicz, Joseph, and Ronald S. Burt. 1991. "Interorganizational Contagion in Corporate Philanthropy." *Administrative Science Quarterly* 36:88-105.
- Gerlach, Michael L. 1992a. *Alliance Capitalism: The Social Organization of Japanese Business*. Berkeley and Los Angeles: University of California Press.
- . 1992b. "The Japanese Corporate Network: A Blockmodel Analysis." *Administrative Science Quarterly* 37:105-39.
- Gerschenkron, Alexander. 1962. *Economic Backwardness in Historical Perspective*. Cambridge, Mass.: Belknap Press.
- Gong Huifeng. 1995. *The [Chinese] Stock Market* (Gupiao Shichang). Shanghai: Shanghai Yuandong Chubanshe.
- Goto, Akira. 1982. "Business Groups in a Market Economy." *European Economic Review* 19:53-70.

- Granovetter, Mark. 1985. "Economic Action and Social Structure. The Problem of Embeddedness." *American Journal of Sociology* 91:481-510.
- . 1995 "Coase Revisited: Business Groups in the Modern Economy." *Industrial and Corporate Change* 4:93-130.
- Greene, William H. 1993. *Econometric Analysis*. New York: Macmillan.
- Hamilton, Gary G., ed. 1991. *Business Networks and Economic Development in East and Southeast Asia*. Hong Kong: University of Hong Kong, Centre of Asian Studies.
- Hamilton, Gary G., and Nicole Woolsey Biggart. 1988. "Market, Culture, and Authority: A Comparative Analysis of Management and Organization in the Far East." *American Journal of Sociology* 94 (suppl.): S52-S94.
- Hamilton, Gary G., and Cheng-shu Kao. 1990. "The Institutional Foundations of Chinese Business: The Family Firm in Taiwan." *Comparative Social Research* 12:135-51.
- Haunschild, Pamela R. 1993. "Interorganizational Imitation: The Impact of Interlocks on Corporate Acquisition Activity." *Administrative Science Quarterly* 38:564-92.
- . 1994. "How Much Is That Company Worth? Interorganizational Relationships, Uncertainty, and Acquisition Premiums." *Administrative Science Quarterly* 39:391-411.
- Jefferson, G. H., and W. Xu. 1991. "The Impact of Reform on Socialist Enterprises in Transition: Structure, Conduct, and Performance in Chinese Industry." *Journal of Comparative Economics* 15:45-64.
- Johnson, Chalmers. 1982. *MITI and the Japanese Miracle: The Growth of Industrial Policy, 1925-1975*. Stanford, Calif.: Stanford University Press.
- Kan Ren. 1996. "Finance Firms Seek Role." *China Daily*, April 29, p. 1.
- Karmel, Solomon M. 1994. "Emerging Securities Markets in China: Capitalism with Chinese Characteristics." *China Quarterly* 140:1105-20.
- Keister, Lisa A. 1998a. "Insider Lending: The Structure and Function of Finance Companies in Chinese Business Groups." In *Organisational Change in Transition Economies*, edited by Daniel Dennison. Ann Arbor: University of Michigan Press, in press.
- . 1998b. "Social Ties and the Formation of Chinese Business Groups." *Sociological Analysis* 1:99-118.
- Kim, Eun Mee. 1991. "The Industrial Organization and Growth of the Korean Chaebol: Integrating Development and Organizational Theories." In *Business Networks and Economic Development in East and Southeast Asia*, edited by Gary G. Hamilton. Hong Kong: University of Hong Kong, Centre of Asian Studies.
- Lamoreaux, Naomi. 1986. "Banks, Kinship, and Economic Development: The New England Case." *Journal of Economic History* 46:647-67.
- . 1991. "Bank Mergers in Late Nineteenth-Century New England: The Contingent Nature of Structural Change." *Journal of Economic History* 3:537-57.
- . 1994. *Insider Lending: Banks, Personal Connections, and Economic Development in Industrial New England*. Cambridge: Cambridge University Press.
- Leff, Nathaniel H. 1978. "Industrial Organization and Entrepreneurship in the Developing Countries: The Economic Groups." *Economic Development and Cultural Change* 78:661-74.
- . 1979. "Entrepreneurship and Economic Development: The Problem Revisited." *Journal of Economic Literature* 17:46-64.
- Li Zhi. 1995. *Modern Chinese Business Groups* (Zhongguo xandai qiye jituan). Beijing: Zhongguo Shangye Ban.
- Lincoln, James R., Michael L. Gerlach, and Christina A. Ahmadian. 1996. "Keiretsu Networks and Corporate Performance in Japan." *American Sociological Review* 61:67-88.

- Lincoln, James R., Michael L. Gerlach, and Peggy Takahashi. 1992. "Keiretsu Networks in the Japanese Economy: A Dyad Analysis of Intercompany Ties." *American Sociological Review* 57:561-85.
- Meeusen, Wim, and Ludo Cuyvers. 1985. "The Interaction between Interlocking Directorships and the Economic Behavior of Companies." In *Networks of Corporate Power*, edited by Frans N. Stokman, Rolf Ziegler, and John Scott. Cambridge: Polity Press.
- Meyer, Marshall W. 1994. "Measuring Performance in Economic Organizations." Pp. 556-78 in *The Handbook of Economic Sociology*, edited by Neil J. Smelser and Richard Swedberg. Princeton, N.J.: Princeton University Press; New York: Russell Sage.
- Miner, Anne S., Terry L. Amburgey, and Timothy M. Stearns. 1990. "Interorganizational Linkages and Population Dynamics Buffering and Transformational Shields." *Administrative Science Quarterly* 35:689-713.
- Miyashita, Kenichi, and David W. Russell. 1994. *Keiretsu: Inside the Hidden Japanese Conglomerates*. New York: McGraw-Hill.
- Mizruchi, Mark S. 1992. *The Structure of Corporate Political Action: Interfirm Relations and Their Consequences*. Cambridge, Mass.: Harvard University Press.
- . 1996. "What Do Interlocks Do? An Analysis, Critique, and Assessment of Research in Interlocking Directorates." *Annual Review of Sociology* 22:271-98.
- Mizruchi, Mark S., and Joseph Galaskiewicz. 1993. "Networks of Interorganizational Relations." *Sociological Methods and Research* 22:46-70.
- Mizruchi, Mark S., and Linda Brewster Stearns. 1988. "A Longitudinal Study of the Formation of Interlocking Directorates." *Administrative Science Quarterly* 33:194-210.
- Munn, C. W. 1981. "Scottish Provincial Banking Companies: An Assessment." *Business History* 23:19-41.
- Naughton, Barry. 1995. *Growing Out of the Plan: Chinese Economic Reform, 1978-1993*. New York: Cambridge University Press.
- Nee, Victor. 1992. "Organizational Dynamics of Market Transition: Hybrid Forms, Property Rights, and Mixed Economy in China." *Administrative Science Quarterly* 37:1-27.
- Oi, Jean C. 1990. "The Fate of the Collective after the Commune." Pp. 15-36 in *Chinese Society on the Eve of Tiananmen: The Impact of Reform*, edited by Deborah Davis and Ezra Vogel. Cambridge, Mass.: Harvard University Press.
- Palmer, Donald, Brad M. Barber, Xueguang Zhou, and Yasemin Soysal. 1995. "The Acquisition of Large U.S. Corporations in the 1960s." *American Sociological Review* 60:469-99.
- Palmer, Donald, Roger Friedland, P. Devereaux Hennings, and Melanie E. Powers. 1987. "The Economics and Politics of Structure: The Multidivisional Form and the Large U.S. Corporation." *Administrative Science Quarterly* 32:25-48.
- Pfeffer, Jeffrey, and Gerald R. Salancik. 1978. *The External Control of Organizations: A Resource Dependence Perspective*. New York: Harper & Row.
- Powell, Walter W. 1990. "Neither Market nor Hierarchy: Network Forms of Organization." *Research in Organizational Behavior* 12:295-336.
- Powell, Walter W., and Peter Brantley. 1992. "Competitive Cooperation in Biotechnology: Learning Through Networks?" in *Networks and Organizations: Structure, Form, and Action*, edited by Nitin Nohria and Robert G. Eccles. Cambridge, Mass.: Harvard Business School Press.
- Powell, Walter W., and Laurel Smith-Doerr. 1994. "Networks and Economic Life." Pp. 368-402 in *The Handbook of Economic Sociology*, edited by Neil J. Smelser and Richard Swedberg. Princeton, N.J.: Princeton University Press; New York: Russell Sage.

- PRC (People's Republic of China). State Council. 1980. *Provisional Regulations Regarding the Promotion of Economic Alliances* (Guanyu tuidong jingji lianhe de zanxing guiding). Beijing: State Council.
- . Central Committee of the Communist Party. 1984. *Regulations of the Central Committee of the Communist Party of China Regarding Economic System Reform* (Zhongguo zhongyang guanyu jingji tizhi gaige de guiding). Beijing: Central Committee of the Communist Party.
- . 1986. *Regulations Regarding a Few Problems Related to the Further Promotion of Lateral Economic Alliances* (Guanyu linyibu tuidong hengxiang jingji lianhe ruogan wenti de guiding). Beijing: State Council.
- . 1987. Committee for Economic System Reform and the National Committee on Economic Alliances. *A Few Ideas Regarding the Establishment and Development of Business Groups* (Guanyu xujian he fazhan qiye jituan de jidian yijian). Beijing: Committee for Economic System Reform and the National Committee on Economic Alliances.
- Reform. 1993. *Zhongguo Jingji Tizhi Gaige Nianjian* (Yearbook of China's economic system reform). Beijing: Gaige.
- Richardson, Jack R. 1987. "Directorship Interlocks and Corporate Profitability." *Administrative Science Quarterly* 32:367–86.
- Rogers, Everett M. 1995. *Diffusion of Innovations*. New York: Free Press.
- Schroath, Frederick W., Michael Y. Hu, and Huiyang Chen. 1993. "Country of Origin Effects of Direct Foreign Investment in the People's Republic of China." *Journal of International Business Studies* 24:277–90.
- Schurmann, Franz. 1965. *Ideology and Organization in Communist China*. Berkeley: University of California Press.
- SASBG (Shanghai Association for the Study of Business Groups). 1995. *The Shanghai Business Groups* (Shanghai qiye jituan fenghua lu). Shanghai: Shanghai Kexue Jishu Chubanshe.
- Shi, Jiyang. 1995. *Modern (Chinese) Banking (Xiandia Yinhang)*. Beijing: Jingji Guanli Chubanshe.
- Spiegel, Mark. 1994. "Gradualism and Chinese Financial Reforms." *Federal Reserve Bank of San Francisco Weekly Letter* 94–44:15–28.
- Steers, Richard M., Yoo Keun Shin, and Gerardo R. Ungson. 1989. *The Chaebol: Korea's New Industrial Might*. New York: Harper & Row.
- Stickney, Clyde P. 1990. *Financial Statement Analysis: A Strategic Perspective*. New York: Harcourt Brace.
- Tilly, Richard. 1966. *Financial Institutions and Industrialization in the Rhineland, 1815–1870*. Madison: University of Wisconsin Press.
- Useem, Michael. 1984. *The Inner Circle*. New York: Oxford University Press.
- Waldner, Andrew G. 1995. "Local Governments as Industrial Corporations: An Organizational Analysis of China's Transitional Economy." *American Journal of Sociology* 101:263–301.
- Williamson, Oliver E. 1981. "The Modern Corporation: Origins, Evolution, Attributes." *Journal of Economic Literature* 19:1537–68.
- . 1985. *The Economic Institutions of Capitalism*. New York: Free Press.
- Xie Cichang. 1996. *The Complete Book of Enterprise Property Rights* (Qiye chanquan shifu quanshu). Beijing: Jingji Ribao Chubanshe.

The Strength of a Weak State: The Rights Revolution and the Rise of Human Resources Management Divisions¹

Frank Dobbin
Princeton University

John R. Sutton
University of California, Santa Barbara

Since the passage of the Civil Rights Act of 1964, federal policy has revolutionized employment rights. Equal employment opportunity law, occupational safety and health legislation, and fringe benefits regulation were designed to create employee rights to equal protection, to health and safety, and to the benefits employers promise. In event-history analyses of data from 279 organizations, this research finds that these legal changes stimulated organizations to create personnel, antidiscrimination, safety, and benefits departments to manage compliance. Yet as institutionalization proceeded, middle managers came to disassociate these new offices from policy and to justify them in purely economic terms, as part of the new human resources management paradigm. This pattern is typical in the United States, where the Constitution symbolizes government rule of industry as illegitimate. It may help to explain the long absence of a theory of the state in organizational analysis and to explain a conundrum noted by state theorists: the federal state is administratively weak but normatively strong.

INTRODUCTION

How does the weak and fragmented U.S. state influence employers? Until the 1970s, sociologists tended to treat organizations as closed systems, immune from external influence. Economists were more attuned to inter-

¹ We thank the National Science Foundation (grant SES-8511250) for support. The first author is indebted to the Copenhagen Business School and to the Danish Social Research Council for providing a research leave. Thanks to John W. Meyer and W. Richard Scott for their collaboration on the larger project this article springs from, to Lauren Edelman and Ann Swidler for their contributions to the pilot study, and to Keith Allum, Bethany Bryson, Jessica Torres, and Roberta Stich for helping us to collect the data. Thanks to Søren Christensen, Julian Dierkes, Paul DiMaggio, Michal Frenkel, Erin Kelly, Donald Light, John W. Meyer, Morten Schmidt, Arthur Stinch-

organizational exchanges but treated government as a transient and usually illegitimate influence on the market. The new organizational institutionalism broke with these traditions by conceptualizing the state as an important force shaping organizations (Meyer and Rowan 1977; DiMaggio and Powell 1983). The rediscovery of the state, encouraged by a renewed emphasis on politics in other areas of sociology (Evans, Rueschemeyer, and Skocpol 1985), has had a salutary corrective effect on organizational research. Recent research shows that, in the United States, several distinct features of the federal government—its structurally limited administrative capacity, dispersal of authority across levels of government, decentralization of decision making at the national level, and ideological deference to the efficiency of the market and the natural virtues of civil society—influence how it regulates organizational life (Meyer 1983; Hamilton and Sutton 1989; Dobbin 1994). The state issues ambiguous mandates to organizations, changes rules frequently in response to protracted political negotiations and litigation, and enforces its rules in a fragmented and indecisive way. Although these features cause it to appear weak, we argue that they produce a peculiar kind of state strength. First, because the state leaves the terms of legal compliance unclear, organizations commit considerable resources to devising compliance measures. Second, because the state signals uncertainty about the legitimacy of its own authority, organizations develop rationales for those compliance solutions that locate authority in the market.

We analyze the effects of the federal employment rights revolution of the early 1970s, which revolved around equal employment opportunity (EEO), health and safety, and benefits initiatives. Previous studies have documented that EEO law was characterized by policy ambiguity and complexity, by a gradual expansion in scope, and by administrative fragmentation (Dobbin et al. 1988; Edelman 1990; Skrentny 1996). We show that these regulatory features also characterized safety and benefits initiatives and that employers responded in like ways to all three initiatives. In all three realms, the state left implementation largely to managers and experts. The uncertainties raised by legal reforms created opportunities for ambitious personnel managers to expand their purview and contributed to the rise of the human resources management (HRM) paradigm. We use retrospective data to show that these initiatives led employers to establish entire new departments dedicated to HRM, EEO, health and safety, and benefits. We show that the diffusion of these departments ac-

combe, John Skrentny, Bruce Western, James Wooten, and the *AJS* reviewers for suggestions. Correspondence may be directed to Frank Dobbin, Department of Sociology, Princeton University, 2-N-2 Green Hall, Princeton, New Jersey 08544-1010. E-mail: dobbin@princeton.edu

celerated after new federal initiatives and that they appeared first in organizations that were vulnerable to governmental scrutiny.

Our central contribution is to document how the U.S. regulatory framework leads managers to recast policy-induced structures in the mold of efficiency. This is the key to the peculiar strength of America's weak state. Middle managers soon retheorized these offices with business and market rationales, describing them not as means to compliance with federal regulations but as cost-effective administrative solutions to problems of worker loyalty and productivity. As a consequence, employers continued to adopt these offices even after Reagan curtailed enforcement of the laws that had popularized them. We suggest that, in the United States, justifications for policy-induced measures drift toward efficiency in this way because constitutional constraints on federal authority symbolize government control of industry as illicit. Compliance measures are thereafter sustained not merely by ephemeral public policies but by transcendental economic laws.

Our argument engages debates about the character of the U.S. state (e.g., Skowronek 1982) and its comparative strength (e.g., Evans et al. 1985). How can the state be administratively weak but normatively strong? How, in other words, can a state with such meager administrative resources shape society so effectively? We argue that the administrative weakness of the state is the cause of its normative strength, for this weakness ensures that Americans will come to see civil society and the market as the sources of social phenomena that are in fact generated by the state. Thus, in the case at hand, constitutional restrictions on federal power that cause Congress to issue complex and ambiguous regulations, forcing employers to devise their own compliance measures, also symbolize federal rule of private enterprise as illicit, encouraging employers to invent efficiency rationales to explain those measures. The Constitution expresses the dangers of state domination of civil society and encourages Americans to see organizational practices and social customs that originate in state action as part of the natural order of things.

After sketching our conceptual argument, we chronicle how U.S. employers responded to new federal initiatives, spelling out hypotheses about where and when they responded by creating new offices. We then review alternative explanations of the creation of offices that emphasize scale economies, unionization, transaction costs, and regulatory effort. Finally, we present multivariate analyses of the diffusion of employment-related offices. We observe 279 employers, many too small to divide personnel functions. Between the mid-1960s and the mid-1980s, personnel/HRM offices doubled in popularity to 70% of the sampled employers, benefits offices doubled to 35%, health and safety offices doubled to over 30%, and equal employment units appeared in 40%.

THE EMPLOYMENT RIGHTS REVOLUTION

Since the early 1960s, much of the elaboration of human rights has taken place in the realm of employment. Congress has sought to extend rights to life and liberty (via health and safety regulation), equal protection (via equal opportunity law), and social citizenship (via regulations insuring fringe benefits) to employees. While these protections are so unevenly institutionalized that they scarcely deserve to be called rights, they nonetheless represent a dramatic change.

None of the legislation encouraged employers to establish new offices, yet, by about 1980, large employers had come to see offices of health and safety, benefits, and equal employment as imperative. These events follow a pattern found in early employment regulation. Ambiguous and complex regulations lead employers to create new departments to manage compliance, and soon, specialists promote these departments as all-purpose solutions to management problems. Conflict over the Wagner Act of 1935 led employers to establish industrial relations (IR) offices and systems, which IR specialists soon promoted for their labor management functions even to nonunion firms (Selznick 1969; Kochan, Katz, and McKersie 1986, p. 45). Complex labor controls during World War II led employers to establish bureaucratic personnel offices and systems, which personnel specialists heralded for their capacity to rationalize hiring and promotion even after wartime labor controls were removed (Baron, Dobbin, and Jennings 1986). Similarly, federal regulations of the early 1970s led employers to establish special compliance offices, which experts then promoted widely for their capacity to optimize human resources and win worker loyalty.

The Character of U.S. Employment Law

Four characteristics of U.S. employment law contributed to the pattern we describe. First, the *ambiguity and complexity of compliance standards* led executives to believe that they would need experts to devise management practices that would comply with the law (Edelman 1992). Ambiguity and complexity result from the weak mandate the Constitution's commerce clause gives Congress to govern private enterprise and the jurisdictional complexity caused by the separation of powers. Second, the *expanding scope* of the law in each realm led executives to believe that they would need permanent offices to track legal changes and devise compliance schemes. The law tends to expand because the presidential system encourages compromise legislation and because autonomous administrative and judicial branches are responsible for interpreting its meaning. Third, the *fragmented nature of regulation* led executives to conclude that organizations would need not merely general compliance offices but special antidiscrimination, benefits, and health and safety offices to (in the

new jargon) "interface" with different regulatory agencies. In the United States, each organization is "likely to have officers that symbolize safety, the environment, affirmative action" to handle interactions with the fragmented bureaucracy (Meyer and Scott 1992, p. 275). Finally, *antistatist elements of the state* encouraged management specialists, intent on justifying their own positions, to retheorize their new offices in efficiency terms and to disassociate them from policy. The Constitution's separation of powers, dispersion of political authority to communities, and limits on federal power symbolize state domination of private enterprise as illicit and inefficient. As Jackall found in studying employer safety practices of the 1970s, "productive return is the only rationale that carries weight within the corporate hierarchy" (1983, p. 58). In consequence, the very journals that in 1975 advised executives to establish specialty offices to pursue legal compliance were by 1985 arguing that these offices helped to rationalize the management of human resources.

Professional groups were at the heart of these developments. Human resources specialists, benefits accountants, tax lawyers, safety engineers, and equal employment managers saw in employment legislation new possibilities for professional growth. They insisted to top managers that new offices could ensure legal compliance, exaggerating the risk of litigation to win organizational resources (see Edelman, Abraham, and Erlanger 1992). In recasting these departments in terms of efficiency, middle managers played to top managers' preoccupation with the bottom line and desire to avoid costs associated with legal compliance.

While we cannot undertake a comparative analysis here, we emphasize that the administratively weak U.S. state differs in significant ways from administratively strong states, among which France is often cited as the paragon. The French state differs on the four characteristics we identified. First, because its constitution does not severely limit state control of private enterprise or fully separate state powers, the French state tends to mandate substantive employment outcomes rather than creating ambiguous and complex regulations (Cummins 1986). For instance, whereas the U.S. state vaguely outlawed discrimination against the handicapped, the French state instructed employers to set aside 3% of jobs for them (Levy 1974). As a consequence, until very recently, French firms had not developed the kinds of internal legal codes of employment that U.S. firms developed (Gavini 1997; Sutton et al. 1994). Second, because the semipresidential system creates a strong president and a unified prime minister, parliament, and administration (a "fusion of powers"), French employment laws are not typically compromise measures that undergo gradual legislative and administrative expansion after passage (Duhamel 1987; Lovecy 1992). For instance, the French state established safety standards, but it has avoided the gradual elaboration of administrative guidelines,

covering hundreds of toxins, that occurred in the United States (Brickman and Jasanoff 1980). Third, because the state is centralized, administration of employment policy is situated in the Ministry of Labor rather than fragmented in a series of autonomous agencies (Meyer 1983). This discourages the fragmentation of the employment function within firms. Fourth, because France's constitution does not contain severe restrictions on government power, it does not symbolize state power as illicit in the way that the U.S. Constitution does (Dobbin 1994). In consequence, French managers face less pressure to recast policy-induced activities in terms of pure efficiency.

ORGANIZATIONAL RESPONSE TO EEO, OSHA, AND ERISA

From the early 1970s, management specialists argued that employers would need offices staffed by experts to comply with new laws. From the early 1980s, they borrowed rhetoric from the emerging human resources management paradigm to argue that these offices improved efficiency generally. Next, we chronicle these events to develop specific hypotheses about the effects of legal shifts, in the process considering the mediating effects of public accountability, organizational age, and employer links to professional groups. In this section, we use John Stuart Mill's (1988 [1974]) method of similarity to show that in three different policy realms, similar regulatory conditions produced similar organizational responses.

The Ambiguity and Complexity of Compliance

Compliance standards in each of these realms were ephemeral, because the law created abstract rights and proscribed various abuses rather than prescribing employer behavior. Civil Rights law established a right to nondiscrimination without specifying how that right was to be protected. Safety and health law established a right to a safe and healthy workplace and defined conditions that threatened employees but did not specify how employers should prevent these conditions. Pension reform law established a right to promised fringe benefits and complex regulations to enforce that right but did not impose a uniform structure for pensions or other benefits. Employers established specialty departments to devise compliance strategies.

Ambiguity in EEO and AA law.—In 1961, John F. Kennedy's Executive Order (EO) 10925 encouraged federal contractors to take "affirmative action" to reverse the effects of past discrimination, and in 1965, Lyndon Johnson's EO 11246 expanded the order and created the Office of Federal Contract Compliance Programs (OFCCP) to enforce it. In 1964, Title VII

of the Civil Rights Act prohibited *all* private employers with 15 or more employees from discriminating on the basis of race, color, religion, sex, or national origin and created the Equal Employment Opportunity Commission (EEOC) to enforce the law (Burstein 1985; Edelman 1990).

Neither Kennedy nor Johnson defined "affirmative action" in their orders to federal contractors, and neither established practical guidelines. Likewise, Title VII made discrimination illegal but did not define it. The act appeared to cover employment practices intentionally designed to discriminate (see Shaeffer 1973, p. 65). Most employers had little idea that the law might apply to them. During the 1960s, the ambiguity of these laws caused only modest concern among employers because sanctions were rare. In the early 1970s, enforcement was increased in both realms, but little was done to clarify the terms of compliance. EEO enforcement was increased by the Equal Employment Opportunity Act of 1972, which extended coverage to very small employers and gave the EEOC power to bring lawsuits itself (Skrentny 1996, p. 127). This came in the wake of the Supreme Court's 1971 *Griggs v. Duke Power* decision, which ruled illegal employment tests that had a "disparate impact" on different groups *absent evidence of intentional discrimination*. The number of Title VII suits skyrocketed, from several hundred a year in the early 1970s to over 5,000 a year in the late 1970s (Burstein and Monaghan 1986). Employers now paid close attention to EEO law (see Petersen 1974).

The OFCCP's Order 4, revised in late 1971, expanded affirmative action coverage to small federal contractors and subcontractors. The majority of U.S. enterprises were now subject to affirmative action law—80% according to a 1973 estimate (Stryker 1996, p. 16). The order required contractors to submit detailed statistical reports and to write annual "affirmative action programs," specifying goals and timetables (Shaeffer 1973, p. 11; Edelman 1992, p. 1537; DuRivage 1985, p. 362). Order 4 required employers to redress inequality but did not clarify how they were to do so beyond stipulating that they could *not* use rigid hiring quotas.

From 1972, employers were expected to invent EEO and AA compliance strategies that would stand up in court, but the courts repeatedly flip-flopped on standards of compliance. As one popular personnel management text warned, "judges interpret [EEO] law differently, and decisions are often reversed in higher courts" (Burack and Smith 1977, p. 183). Management journals extolled voluntary hiring quotas, written job tests to ensure objective selection of employees, and bureaucratic hiring and promotion systems to undermine cronyism (Dobbin et al. 1993), but the courts proved unpredictable, vacillating on the legality of hiring quotas, employment tests, and a variety of other measures. So variable was the law that the Bureau of National Affairs published the Equal Employment

Opportunity Manual in the form of a serial (Ropp 1987). Managers devised an astonishing diversity of antidiscrimination practices—by the end of the 1970s, Kenneth Marino found 33 new practices in wide use (1980).

The continuing ambiguity of compliance standards led management writers to advocate permanent antidiscrimination offices to track legal shifts. Because the courts were so fickle, Marino (1980, p. 25) advised executives to adopt the "Good-Faith-Effort Strategy," the heart of which was a special office designed to signal that the employer was making every effort to figure out how to comply. In a *Harvard Business Review* article, Antonia Chayes (1974, p. 81) noted that "vigorous enforcement" had brought "serious top management attention to antidiscrimination legislation. . . . Now the penalties imposed under employment discrimination laws are seen as posing a severe financial threat." She advised executives to set up EEO and AA programs that could prevent lawsuits. Meanwhile, compensation of upper managers was being tied to affirmative action performance, and this led them to support dedicated antidiscrimination departments (Vernon-Gerstenfeld and Burke 1985, pp. 59–60; Johns and Moser 1989; Ropp 1987).

Indefinite safety and health regulations.—The Occupational Safety and Health Act of 1970 represented a revolution of sorts, for it required all employers to take steps to prevent illness and injury rather than merely making restitution after the fact through Workmen's Compensation. The act created the Occupational Safety and Health Administration (OSHA), with the power to codify safety and health regulations as they emerged in legislation, case law, and administrative law. It gave states the option of using OSHA inspectors or state safety inspectors to monitor workplaces and respond to complaints (Bardach and Kagan 1982, p. 46).

Administrative health and safety guidelines outlawed specific workplace conditions in excruciating detail rather than prescribing how to prevent those conditions, in part because the Constitution checked Congress's capacity to dictate to private enterprises. OSHA wrote countless guidelines that made particular conditions illegal but left local safety engineers to figure out how to comply. In the case of health hazards, OSHA established guidelines for exposure to hundreds of toxins. It specified acceptable levels of asbestos particulates, for instance, without dictating how factories should achieve those levels—whether by changing production techniques, installing filters, or automating work. In the case of safety hazards, OSHA established broad guidelines and encouraged plant-specific solutions. Guides to compliance, such as Richard Anderson's (1975) *OSHA and Accident Control through Training*, suggested that employers set up safety departments, training programs, safety committees, grievance mechanisms, and "practice" federal safety inspections (Foulkes and Morgan 1977; Ritter and Wagel 1989).

As John Meyer and Brian Rowan argued, government safety controls “make it important for organizations to create formal safety rules, safety departments, and safety programs” to symbolize their commitment (1977, p. 350). The personnel journals suggested that such offices served a public relations function, contributing to “the appearance of social responsibility” (Mendelhoff 1979, p. 92), and could help inoculate firms against damaging legal judgments by signaling that the employer was making a “good faith effort” to comply with complex and changing regulations (Anderson 1975; Ewing 1983).

Ambiguity and complexity in benefits regulation.—The Employee Retirement Income Security Act (ERISA) of 1974 was designed to guarantee that employers would make good on pension and benefit promises. ERISA also used tax incentives to encourage employers to offer benefits democratically, to all employees. It regulated not only common retirement schemes but the whole range of employee benefits (Gill 1985, pp. 2–8). The act was a maze of esoteric legal language and was widely viewed as “the most complex piece of legislation ever passed by Congress” (Tepper 1977, p. 105). It did not dictate a universal form for pensions or other benefits, but it regulated them in fine detail. As amended through 1984, the act and its Internal Revenue Code amendments ran to 320 pages (Gill 1985). For both pension plans and benefit plans, the act established (a) federal reporting requirements, (b) requirements for disclosure to employees, (c) fiduciary requirements, and (d) investment regulations. For pension plans, the act also established (e) pension plan insurance, (f) eligibility and participation rules, (g) vesting guidelines, and (h) funding regulations (Coleman 1985).

Despite the tremendous complexity of the legislation, key stipulations remained ambiguous. The act demanded that pension plan managers exercise “fiduciary responsibility” but offered neither a definition nor a litmus test. As one primer on benefit regulation warned: “no one knows what the judicial interpretation” may be (Logue 1979, p. 65). The act required that pensions be backed by certain levels of investment, without specifying precisely how funds were to be invested. It gave employers three different sets of pension vesting guidelines to choose from. Compliance was not simply a matter of following a blueprint—it required accountants and tax attorneys who could weigh a dizzying mix of options.

While compliance standards were complex and ambiguous, the costs of noncompliance were high. Rather than requiring employers to comply, as safety and health legislation did, the act created financial incentives. First, for complying plans, employer expenses are tax deductible, employer contributions are not taxed as current income to employees, and earnings on invested funds are not taxed (Klein 1986, p. 72; Skolnik 1976). Second, the act guaranteed pensions both through insurance *and with the assets of the employer*, permitting employees to make claims not only

against pension funds but also against employers. This meant that employers bore the risk of subscribing to poorly managed plans (Logue 1979, p. 62; Wooten 1994).

Employers had to rewrite virtually all pension plans and many other fringe benefit plans to comply with the new regulations (Hakala and Huggins 1976; Meyer and Fox 1974). One unanticipated consequence of the act was that thousands of small companies chose to abandon their pension plans rather than bear the costs of compliance (Davey 1978, p. 4; see also Achenbaum 1986, p. 149; Logue 1979, p. 68). Most of the rest rewrote their plans—at least 300,000 companies did so immediately (Klein and Moses 1974).

The personnel journals counseled executives to consult tax lawyers and accountants in writing plans and to establish departments to monitor plans and prepare reports for the government. A new breed of “benefits consultants” emerged to argue that experts alone could devise schemes that would guarantee tax deductions, fend off litigation, and minimize employer outlays (Meyer and Fox 1974, p. 53; Meyer 1981).

The Expanding Scope of the Law

In each realm, the scope of the law was expanded frequently via new legislation, administrative rulings, and case law. As a consequence, personnel specialists, antidiscrimination experts, safety engineers, and benefits accountants lobbied for special departments to track changes in the law and create new compliance measures.

New classes of discrimination.—Congress first expanded the scope of the Civil Rights Act in 1967, adding employees over 40 to the list of protected groups (Farley 1979, p. 12). Following the Equal Employment Opportunity Act of 1972, which signaled that the EEOC would enforce EEO law with vigor, Washington expanded the scope of the law often. In 1973, Congress required federal contractors to extend affirmative action to the handicapped, and in 1974, it required them to include Vietnam-era veterans (Pati and Adkins 1980).

Administrative and judicial rulings were responsible for much of the growth in the law. For instance, neither sexual harassment nor maternity leave were initially covered under sex discrimination law, but during the 1970s, both the EEOC and the courts came to treat harassment and refusal to grant maternity leave as sex discrimination (Bradshaw 1987, p. 51). Sexual harassment guidelines illustrate how management specialists responded. In 1980, the EEOC issued formal guidelines that read, “Harassment on the basis of sex is a violation of Section 703 of Title VII,” but offered no hints about how to prevent harassment (quoted in Spann 1990, p. 58). Key management journals—*Personnel*, *Public Personnel*

Management, the *Harvard Business Review*—warned of the potential for litigation and advised managers to establish a “good faith effort” to prevent harassment by creating training programs for supervisors, hiring ombudspersons, establishing grievance mechanisms, and publishing guidelines (Spann 1990). They advised executives to establish equal employment opportunity/affirmative action (EEO/AA) offices to undertake these measures.

OSHA's growing list of dangers.—The scope of OSHA regulation expanded dramatically as a result of administrative rulings. By the late 1970s, OSHA had developed standards for some 400 dangerous substances, covering the issues of air sampling, personal protection, handling procedures, and medical exams (Northrup, Rowan, and Perry 1978, p. 49). As OSHA regulations grew, management writers implored firms to hire safety specialists. In *Regulating Safety*, John Mendelhoff (1979) argued that OSHA regulations had become so complex that firms could not depend on line managers for compliance. As early as 1978, a study of the chemicals and aerospace industries found that the act had caused the typical firm to create a safety department or to enlarge the existing department (Northrup et al. 1978, p. 224).

It was the instability of the law that made a safety department imperative. As Burton Malkiel argued in 1979 in the *Harvard Business Review*, “changing health and safety regulations” undermine corporate planning: “It is not so much the direct cost of regulation that has inhibited investment and R&D but rather the unpredictability of regulatory changes” (1979, p. 90). Administrative and judicial decisions constantly altered the rules of the game.

The expansion of benefits regulation.—After 1974, Congress altered the tax code and amended ERISA a number of times, complicating the matter of compliance and expanding the scope of the law. Details of the tax treatment of various retirement plans were amended in new tax code legislation in 1975, 1978, 1981, 1982, and 1984 (Coleman 1985, p. 10). Each change required employers to review and rewrite their pension plans (Meyer 1981; Bixby 1986). The biggest change to ERISA came in 1980, when Congress expanded regulation of small employers, making those in shared pension plans liable for up to 100% of their assets and spurring even tiny firms to hire experts to manage pension programs (Figgie 1981).

The scope of ERISA coverage also grew for individual employers, as they switched from one-size-fits-all benefits programs to myriad pension and health insurance options and as they added such benefits as permanent disability insurance and on-site daycare. The management journals advised firms to hire full-time specialists who could design benefit plans to comply with the law and to anticipate future legal changes. For instance, as employers added 401k plans, benefits accountants advised them

to make low-wage workers eligible, so as to ensure tax-exempt status, and to sell the plans to low-wage workers, in case the courts should decide to treat plans that discriminate de facto as taxable (Moody and Higgins 1984).

The Fragmentation of Enforcement

Enforcement of these laws was carried out by autonomous federal agencies with distinct charges. Management writers argued that all-purpose, employment-law experts would not suffice in America's fragmented regulatory environment. A 1977 article in the *Harvard Business Review* advised: "The various requirements of state and federal regulations . . . make increasing demands on both profit and nonprofit organizations. . . . Compliance with the laws relating to OSHA, EEOC, and ERISA demands expertise" (Foulkes and Morgan, p. 160).

Interfacing with the EEOC and OFCCP.—The EEOC and OFCCP were established as independent agencies to enforce EEO and AA law respectively. While nothing in the law required employers to establish antidiscrimination offices, from the early 1970s, the management journals counseled them to do so in articles with titles such as "A Total Approach to EEO Compliance" (Giblin and Ornati 1974, p. 37). Special offices could create antidiscrimination programs to preclude litigation by the EEOC and could deal with OFCCP demands for employment statistics, written affirmative action goals and timetables, and on-site "compliance reviews."

Barbara Boyle (1973, pp. 88–89) wrote in the *Harvard Business Review* that costly litigation made an autonomous antidiscrimination office "at the highest practicable level in the organization (i.e., outside the personnel office)" well worth the price: "The establishment of an affirmative action program is not costly—its absence is." Antonia Chayes (1974) followed with an article in the same journal, titled "Make Your EEO Program Court-Proof," that recommended naming an affirmative action officer at a minimum. A 1974 article in *Personnel* suggested that the law provided an opportunity to establish personnel departments and full-blown personnel systems: "Viewing the [antidiscrimination] guidelines in strictly negative terms, executives have failed to see that the government concern for programs of EEO compliance actually provides the impetus for developing a personnel system" (Froehlich and Hawyer 1974, p. 62). Employers responded by creating EEO and AA departments (Giblin and Ornati 1974, p. 45; Johns and Moser 1989, p. 56).

Interfacing with OSHA.—OSHA's powers extended beyond data collection and litigation to inspection of workplaces. The threat of inspection loomed large because OSHA regulations were already daunting and were

proliferating quickly. As Herbert Froelich and Dennis Hawyer counseled in the journal *Personnel* in 1974 (p. 67): "Some years ago, a vice-president of General Motors said that 'safety is good business.' He may not have anticipated the tough regulatory crackdown of OSHA, but there is a parallel in the current government regulations: Compliance is good business." Compliance manuals suggested that employers have safety experts on hand: "An excellent person to guide the [OSHA] inspector would be the safety manager or safety engineer" (Anderson 1975, p. 243). The engineer might direct the inspector's attention to major safety innovations and away from minor technical violations (Rees 1988, p. 53). Rees found that by the mid-1980s, executives had hired safety experts with the specific goals of tracking case law and educating OSHA inspectors (1988, p. 52).

From the early 1970s, engineering journals, personnel journals, and management treatises advised that safety experts could now win permanent positions and even their own departments. The authors of a study conducted in the mid-1970s concluded: "There was general agreement among safety personnel in a broad range of firms that the passage of the Act had served to increase their own status and credibility and that of their function within the company" (Northrup et al. 1978, p. 224). The act had boosted the importance of safety units: "The net effect has been a substantial increase in the status of health and safety personnel. . . . They are regarded as necessary defenders against costly and embarrassing mistakes that threaten corporate viability" (Northrup et al. 1978, p. 49).

Interfacing with the IRS and the PBGC.—Enforcement of ERISA was divided between the Treasury Department's Internal Revenue Service (IRS), which held the authority for tax disqualification and which governed most funding issues, and the Labor Department's Pension Benefit Guarantee Corporation (PBGC), which governed fiduciary issues and handled reporting and disclosure. The PBGC demanded a range of different reports from employers, including an annual disclosure statement for each employee (Klein 1986; Coleman 1985, p. 7).

Writers in insurance and management journals argued that the paperwork alone augured well for the creation of personnel and benefits offices. ERISA and Social Security changes, a 1980 article in *Personnel* concluded, had boosted executive support: "Personnel activities in many organizations are becoming increasingly important—largely because of the high cost of labor and benefits [and] the negative impact of government regulations" (Zippo 1980, p. 66). As a 1977 article in the *Harvard Business Review* concluded, "Personnel must have the ability to understand the needs both of the line organization and of the various branches of state and federal government. . . . We suggest that a team be developed" that includes specialists in benefits and in the law (Foulkes and Morgan

1977, p. 162). By failing to adhere to ERISA's complex pension guidelines, the journals pointed out, employers risked losing tax deductions for contributions, having their plans terminated by the PBGC, and being sued by current and former employees. The journals advised firms to hire benefits experts with accounting and legal backgrounds who could document compliance and make sense of emerging legal standards.

In sum, in each realm, new employment laws were ambiguous and complex, were subject to frequent expansion, and were enforced by distinct federal agencies. We hypothesize that executives responded to these laws by creating specialized departments. The journals advised employers to establish personnel offices as a first step, thus, we expect personnel offices to rise from the time of the first legislative change.

HYPOTHESIS 1.—*The Equal Employment Opportunity Act of 1972, coupled with the OFCCP's Revised Order 4 from late 1971, caused the adoption of EEO/AA offices to rise from 1972.*

HYPOTHESIS 2.—*Passage of the Occupational Safety and Health Act caused the adoption of health and safety offices to rise from 1970.*

HYPOTHESIS 3.—*The Pension Reform Act caused the adoption of benefits offices to rise from 1974.*

HYPOTHESIS 4.—*Equal employment, safety and health, and pension reform legislation caused the adoption of personnel offices to rise from 1970, when the first legislative landmark occurred.*

We argue that public accountability and age should mediate the effects of legal shifts.

Public accountability and the symbolization of justice.—W. Richard Scott and John Meyer (1987) argue that public and nonprofit organizations are first to embrace new norms because they are judged more by their activities than by their performance. Marshall Meyer (1979, p. 285) adds that public organizations can more readily pass on the costs of new structures. We expect that public and nonprofit organizations will be more likely to install new offices despite the fact that government offices were exempt from many of these regulations.

HYPOTHESIS 5.—*Public and nonprofit organizations adopted new rights-related offices more readily than did private organizations.*

Organizational age and institutionalization.—Stinchcombe (1965) argues that managers adopt the organizational structures prevalent at the time of the industry's birth. New organizations and industries are most likely to have the latest practices as a consequence. Studies have confirmed this and have shown that older organizations are more highly institutionalized in Selznick's (1949) sense: they are resistant to change.

HYPOTHESIS 6.—*Older organizations were less likely to adopt each type of office.*

Antistatist Elements and the Reconstruction of Offices as Efficiency Centers

Middle managers initially promoted equal employment, health and safety, and benefits departments to manage legal compliance. During the late 1970s and early 1980s, however, they reconstructed these departments as components of human resources management and articulated theories of how each contributed to productive efficiency. Managers did this, we argue, because the Constitution paints state domination of private enterprise as illicit. By the time the Reagan administration cut enforcement in these three realms, the new justifications had succeeded so well that employers continued to create new departments.

Here our predictions contrast sharply with those of legal analysts who note that Reagan cut federal regulation with much fanfare and who thus predict a weakening of organizational compliance. Students of law and regulation, discussed below, suggest that compliance efforts will be proportional to the threat of regulatory sanction. Rather than declining in popularity during the early 1980s, we expect these offices to increase in popularity.

The reconstruction of the EEO/AA office as the "diversity" center.—The renovation of antidiscrimination offices as efficiency centers began almost immediately. First, personnel specialists argued that formal hiring and promotion rules inspired by EEO law would not only help employers fight discrimination suits but would help them to rationalize the placement of employees. As early as 1974, Froelich and Hawyer (pp. 62–63) argued in *Personnel* that equal employment opportunity law had spawned performance-based personnel systems, which should be "as much a component of sound business planning as financial, manufacturing, and market planning are." By 1979, a *Wall Street Journal* poll of top executives found that nearly two-thirds favored government programs to increase hiring of women and minorities (*Harvard Law Review* 1989, p. 661). Early in the 1980s, management specialists described affirmative action practices as an "essential management tool which reinforces accountability and maximizes the utilization of the talents of [the] entire work force" (Feild 1984, p. 17). When Reagan sought to end enforcement, he encountered substantial opposition from business leaders on just these grounds (*Harvard Law Review* 1989).

Second, personnel specialists came to argue that diversity in the workplace increases efficiency in and of itself. From the late 1970s, management consultants extolled the secondary benefits of workforce diversity. "Leading consultants, academics and business leaders . . . point out that a well managed, diverse workforce holds potential competitive advan-

tages for organizations" (Cox and Blake 1991, p. 45). In human resources parlance, "diversity" came to replace "affirmative action" as the code word for efforts to integrate the workforce. Proponents argued that diversity offers design, production, and marketing advantages because it brings people with varied skills and backgrounds together. In 1986, the National Association of Manufacturers described affirmative action as a "business policy which has allowed industry to benefit from new ideas, opinions and perspectives generated by greater workforce diversity" (quoted in *Harvard Law Review* 1989, p. 669; see also Bureau of National Affairs 1986, p. 93). We expect, following Edelman (1992), that despite Reagan's cuts in EEOC enforcement, diffusion of personnel and EEO/AA offices continued apace.

The reconstruction of health and safety as HRM and publicity functions.—Middle managers transformed OSHA-inspired safety programs into efficiency measures in three ways. First, they argued that new safety initiatives demonstrated the firm's commitment to employees, the key to winning employee commitment to the firm according to the HRM paradigm. As a popular HRM text suggested as early as 1977: "A manager can use health and safety to motivate workers" by decreasing "the sense of alienation" and increasing "identification with the organization's purposes" (Burack and Smith 1977, p. 419).

Second, during the 1970s and 1980s, executive remuneration came to be linked to stock price, causing top managers to become acutely sensitive to negative publicity. In 1983, Robert Jackall reported in the *Harvard Business Review* that executives believed that "the bad publicity from one serious accident in the workplace can jeopardize years of work and scores of safety awards." In the words of one high-ranking chemical company executive: "In the corporate world, 1,000 'Attaboys!' are wiped away by one 'Oh, shit!'" (Jackall 1983, p. 128). Engineers thus sold health and safety departments to executives to prevent accidents that bring negative publicity, threatening stock price, market share, and executive compensation.

Third, managers argued that new safety technologies spurred them to replace antiquated production facilities and gave them a first-mover advantage in a global economy with a growing market for safe and environmentally sound technologies. The response to OSHA's 1978 lowering of the permissible level of airborne cotton dust illustrates this point. The textile industry fought cotton dust controls designed to end brown lung disease, but the Supreme Court eventually upheld OSHA's ruling. Meanwhile, managers at one company automated cotton manufacture to reduce cotton dust exposure. They sold the changes to executives by documenting productivity gains, arguing that "OSHA's regulation on cotton dust has

been the main factor in forcing technological innovation in a centuries-old and somewhat stagnant industry" (Jackall 1983, p. 50).

The reconstruction of benefits as the new corporate "welfare work."—In the 1920s, many employers joined the "welfare work" bandwagon, convinced by arguments that generous fringe benefits could win employee commitment and undermine unionism (Brandes 1976). From the late 1970s, parallel arguments for corporate welfarism were proffered by benefits and management specialists. First, they argued that the key to winning employee commitment and solving recruitment and retention problems was to display the employer's commitment to the employee with pension, health, training, and family programs. Second, they advised that well-conceived benefit plans benchmarked to those of competitors could stem union efforts (Foulkes 1980, p. 227).

Experts clothed the new welfarism in the rhetoric of modern management, emphasizing the importance of strategic planning and integration with other management functions. One benefits textbook marveled that, "prior to the early 1980s, strategic planning [of compensation and benefits] was rare" (McCaffery 1986, p. 18). Robert Greene and Russell Roberts wrote in 1983 that sophisticated benefits packages should take employer resources into account, with an eye to "what will be required to meet the strategic goals of the organization" (p. 82). In Fred Foulkes's (1980) study of personnel and benefits systems, managers cited competitive advantage in the labor market, not legal compliance, as the key function of benefits departments. Consultants argued that a professional department charged with maximizing benefits while minimizing cost can save the employer huge sums on turnover costs, antiunion efforts, and benefits packages themselves.

We expect that despite cutbacks in federal enforcement in these three realms during the Reagan administration, which we detail below, the reconstruction of new departments in efficiency terms caused them to diffuse even after 1980.

HYPOTHESIS 7.—Despite the reduction in enforcement after 1980, the reconstruction of EEO/AA, health and safety, and benefits offices in efficiency terms caused adoption to remain high. Despite the reduction in enforcement, the rise of the human resources management paradigm caused adoption of personnel/HRM offices to remain high.

The Stages-of-Institutionalization Thesis

The reconstruction of policy-induced structures as efficient should not only cause diffusion to continue after the Reagan administration cut enforcement but should produce the pattern that Tolbert and Zucker (1983)

discovered in the case of early civil service reforms. First, the structure is prescribed as a solution to a particular problem, whether it be a functional problem (e.g., hazardous work) or a problem of legal compliance. Next, the structure is prescribed as part of accepted management practice; "The legitimacy of the procedures themselves serves as the impetus for the late adopters" (Tolbert and Zucker 1983, p. 35). In consequence, the organizational factors that predict early adoption do not predict late adoption. We expect to find that the organizational characteristics that predict adoption of personnel, benefits, health and safety, and antidiscrimination offices will decline in importance over time, as these offices are transformed from compliance centers into productivity centers.

HYPOTHESIS 8.—*The institutionalization of personnel, EEO/AA, health and safety, and benefits units caused characteristics predicting adoption to decline in importance over time.*

Professionals as Agents of Diffusion

Professional groups theorized the importance of these offices (Strang and Meyer 1994; Sutton and Dobbin 1996). Three groups were particularly important: personnel professionals, labor relations specialists, and lawyers. We argue that even loose links to such professionals should affect adoption of offices.

Personnel, labor relations, and legal offices.—A personnel office is frequently the first line of defense executives establish against employment litigation. While the personnel journals frequently advocated the creation of separate EEO, health and safety, and benefits offices as further defenses, previous studies have shown that when other factors are controlled, personnel offices are *negatively* associated with the creation of antidiscrimination offices (Dobbin et al. 1988; Edelman 1992). Personnel managers apparently prefer to take on new functions themselves.

HYPOTHESIS 9.—*Organizations with personnel offices were less likely than others to adopt benefits, health and safety, and EEO/AA offices after legal landmarks.*

Labor relations specialists, by contrast, make organizations sensitive to employment law but do not compete for resources with antidiscrimination, benefits, or health and safety offices. We expect that the presence of a labor relations unit will have a positive effect on these offices.

HYPOTHESIS 10.—*Organizations with labor relations offices were more likely to adopt benefits, health and safety, and EEO/AA offices after legal landmarks.*

Legal offices also make organizations sensitive to employment law (Sut-

ton and Dobbin 1996), and like labor relations offices, they do not compete with the other offices.

HYPOTHESIS 11.—*Organizations with legal offices were more likely to adopt benefits, health and safety, and EEO/AA offices after legal landmarks.*

Weaker links to the personnel and legal professionals.—As we have seen, the journals of personnel associations, such as *Personnel* and *Personnel Journal*, actively promoted the creation of new personnel, equal employment, benefits, and health and safety positions. Other studies have found that personnel association membership increases the likelihood that an organization will adopt legal compliance measures (Dobbin et al. 1993). We expect this to hold true for HRM divisions.

HYPOTHESIS 12.—*Membership in a personnel association encouraged adoption of personnel, benefits, health and safety, and EEO/AA offices after legal landmarks.*

Links to labor and employment-law attorneys should have a particularly strong effect on the creation of compliance departments. We asked employers when, if ever, they kept a labor or employment-law attorney on retainer.

HYPOTHESIS 13.—*Organizations keeping a labor or employment-law attorney on retainer were more likely to create personnel, benefits, health and safety, and EEO/AA offices after legal landmarks.*

Labor Relations Law

We have argued that several offices were popularized by the rights revolution of the early 1970s. But it is also possible that personnel-related offices diffused as part of a secular trend. To rule this out, we analyze labor relations offices. Between 1955 and 1985, there were no major legal reforms that might have increased their popularity. The most important federal legislation, the Landrum-Griffin Act of 1959, regulated union finances and strike activity rather than expanding union rights. The most important state legislation, "right to work" laws adopted by 20 states by the end of the 1970s, actually curtailed union rights.

HYPOTHESIS 14.—*In the absence of landmark labor legislation, the adoption of labor relations departments did not rise significantly during the 1970s.*

Two important changes in labor relations did occur, however. First, there was a general decline in unionization. Because this change occurred largely through the founding of nonunion firms in new industries, we do not expect to find that employers closed existing labor relations offices. Second, unionization grew among white collar employees in the govern-

ment and nonprofit sectors during the 1970s. We expect that this helped to institutionalize labor relations offices. Thus, we expect to see an increase in labor relations departments in those sectors, even statistically controlling for whether the employer is unionized in the current year.

HYPOTHESIS 15.—*Due to growing government and nonprofit unionization, the adoption of labor relations departments rose in those sectors during the 1970s.*

Next, we review competing theories of office creation, which point to scale economies, unionism, transaction costs or labor segmentation, and regulatory effort. In the section that follows, we explore hypotheses with event-history analyses of data collected from 279 organizations for the period 1955–85.

ALTERNATIVE THEORIES OF OFFICE CREATION

Several theories offer hypotheses about where and when personnel, benefits, health and safety, equal employment, and labor relations offices will appear.

Scale Economies

In “A Formal Theory of Differentiation in Organizations,” Peter Blau (1970) argues that increases in scale are accompanied by the proliferation of subunits. Blau sees increases in size leading to specialization within the organization, as in Adam Smith’s celebrated pin factory, and consequently to differentiation. Blau implies that personnel administration will be differentiated from general management once the organization reaches a certain size and that functions such as benefits administration will be differentiated once the organization reaches another threshold (see also Blau and Schoenherr 1971). Empirical studies have shown a declining effect of size, represented by a logged measure.

HYPOTHESIS 16.—*Large organizations were more likely to install all five types of offices.*

Unionism

Safety and fringe benefits are two of the pillars of union politics in the United States. Early studies of personnel practice show that unions caused employers to create new offices to deal with these issues (National Industrial Conference Board 1940). More generally, from the time of the Wagner Act, unionized firms have installed both personnel and labor relations departments to stem the growing power of unions (Jacoby 1985).

HYPOTHESIS 17.—*Unionized organizations were more likely to adopt health and safety, benefits, personnel, and labor relations offices.*

Transaction Costs and Labor Segmentation Theories

Transaction costs and labor segmentation theorists find that firms in core, capital-intensive sectors go to great lengths to secure long-term employment (Williamson 1975; Gordon, Edwards, and Reich 1982; see also Doeringer and Piore 1971). They use internal promotion systems, seniority rules, and benefits linked to job tenure, establishing personnel and benefits offices to administer these programs. They establish labor relations offices to handle unions (Jacoby 1985). They should also create safety offices to prevent accidents and equal employment offices to prevent disaffection of women and minorities. Among the 13 sectors in our sample, chemicals and electrical manufacturing are most easily classified as core. Machinery is a mixed bag, containing core firms and small machine shops (Gordon et al. 1992, pp. 199–201), but our sample probably contains few machine shops because it is limited to firms with at least 50 employees and \$500,000 in assets as of 1985. As early as the 1940s, a similar sample of machinery firms showed a high incidence of personnel departments (Baron et al. 1986).

HYPOTHESIS 18.—*Organizations in the chemicals, machinery, and electrical industries were more likely to adopt all five kinds of offices.*

Regulatory Effort

Students of law and regulation from both left and right agree that active enforcement and the threat of sanctions increase corporate compliance activities (Sunstein 1996; Posner 1997). Enforcement of employment law was most active during the 1970s. Ronald Reagan ran for president on a promise to reduce federal regulation of business and made good on that promise soon after taking office in 1981. Reagan officials directed the Equal Employment Opportunity Commission (EEOC) to curtail litigation and loosen standards and halved staffing at affirmative action's oversight agency, the OFCCP (Edelman 1992, p. 1541; Skrentny 1996; DuRivage 1985, p. 364). Reagan cut PBGC and Department of Labor enforcement of the Pension Reform Act, and, by cutting antitrust enforcement, he enabled firms to acquire others for the purpose of raiding their pension schemes (Achenbaum 1986, pp. 155–56). Reagan directed OSHA to curtail inspection of workplaces, to halt ongoing litigation against employers, and to reduce the number of regulated toxins (Jackall 1983; Ewing 1983; Hartnett 1996). Perhaps most important, Reagan's judicial and quasi-judicial

appointees opposed workplace regulation and thus weakened enforcement efforts.

HYPOTHESIS 19.—*Employers were more likely to adopt EEO/AA, benefits, health and safety, and personnel offices in the 1970s when regulatory efforts were greatest, and they were less likely to adopt these offices from 1981, when regulatory efforts declined.*

DATA AND METHODS

In the previous section, we chronicled the history of policy in three realms to show that the federal government's policy style produced a similar chain of events in each. New policies generated specialty compliance departments. Antistatist sentiments led incumbents in those departments to develop efficiency justifications for their existence. In consequence, the withdrawal of federal enforcement did not stall the diffusion of those departments. And because those departments had been theorized as enhancing efficiency in general, the factors associated with early adoption became less important over time. Next we test these claims with data from several hundred employers.

The Sample

We selected a stratified random sample of public, for-profit, and nonprofit organizations in 1985, collecting retrospective data from 279 organizations. We generated the sample in three states with different legal environments: California, New Jersey, and Virginia (Sutton et al. 1994). We sampled from 13 diverse sectors to make tests of sector effects possible.

The difficulties associated with sampling organizations have been well documented (Kalleberg et al. 1990). We sampled randomly from the best published list of organizations available for each sector. For firms in banking, chemicals, electrical manufacturing, machinery, retail trade, and transportation, we went to *Dun's Million Dollar Directory*, which lists all publicly traded firms with assets of \$500,000. We sampled hospitals from the directory of the American Hospital Association (1983) and nonprofits from the *Encyclopedia of Associations* (Akey 1984). For public organizations, we used state and federal agency directories and telephone directories to sample equal numbers of city, county, state, and federal employers. Before identifying a final sample, we contacted employers to ensure that they were located in one of the three states, operated in one of the 13 sectors, and employed at least 50 persons.

Response rate.—We contacted 620 organizations and received completed questionnaires from exactly 300, for a response rate of 48%. After follow-up telephone calls to fill in missing data, we excluded 21 organiza-

tions due to poor data quality. This brought the survey completion rate down to 45%, which compares favorably with other organizational studies: Blau and colleagues (1976) completed 36%, Lincoln and Kalleberg (1985) completed 35%, and Edelman (1992) completed 54% in a telephone interview. The sample may suffer from selection bias, because organizations with personnel offices were probably more likely to respond. However, the over-time framework minimizes the problem by ensuring many cases (annual organizational spells) without personnel departments, particularly in the early years.

Measurement

Dependent variables.—We examine the creation of five structures: personnel/HRM, benefits, health and safety, EEO/AA, and labor relations. We do not distinguish between personnel and HRM departments, because they are functional equivalents. In the cases of personnel/HRM, benefits, health and safety, and labor relations, we analyze the establishment of offices.

In the case of EEO/AA, we analyze creation of either an office *or* an officer for several reasons. First, management journals often advocated affirmative action “officers”; whereas, they advocated benefits and health/safety “offices” or “departments” (see Chayes 1974). Second, Edelman (1992) found that the average size of corporate affirmative action *offices was less than two employees*. Thus, the difference between an “office” and an “officer” is often semantic—either may involve a single expert. Third, because the average affirmative action office in a public agency has over seven employees (see Edelman 1992), we found that public employers tend to use “office” and private employers tend to use “officer.”

Our data cover 1955–85. For EEO/AA office or officer, we begin the analysis in 1961, the year in which John Kennedy required “affirmative action.” For other offices, which existed in some organizations before 1955, we begin the analysis in that year. For all outcomes, organizations born after 1955 enter the risk set upon birth. Organizations that abandon a particular office reenter the risk set. None of these offices was abandoned by more than a handful of employers.

Independent variables.—Table 1 lists the independent variables. All vary over time. We expect the Occupational Safety and Health Act of 1970, the Pension Reform Act of 1974, and the EEO Act of 1972 to increase the adoption of related departments. We expect these three changes to affect personnel/HRM departments from 1970, and we examine whether the diffusion of labor relations offices rises in the same year.

For each department, we start the third time period in 1981 to examine whether Reagan’s cuts in enforcement stalled diffusion. If we are correct

TABLE 1
VARIABLE LIST

Variable	Definition
Log employment	Natural logarithm of number of employees
Union	Presence of a union contract
Personnel/HRM office ..	Presence of a personnel or human resources management office
Labor relations office	Presence of a labor relations office
Personnel association ..	Member of a personnel association
Age of organization	Age of organization in years
Periods	Periods of relevant policy regimes (see text)
Legal office	Presence of a legal office
Labor attorney on retainer ..	Organization has a labor/employment attorney on retainer
Sectors *	
Electrical and machinery	
Chemicals	
Local government	Includes local and county
State and federal	
Nonprofit	

NOTE—All variables are measured at the start of the year. All vary over time.

* Omitted categories: printing, transportation, trade, banking, hospitals.

in arguing that these offices had been justified in purely economic terms, the rate of adoption should not decline.

We include binary variables representing each of the sectors of theoretical interest. To operationalize transaction costs and labor segmentation theories, we include capital intensive industries with firm-specific skills: chemicals, machinery, and electrical machinery. To operationalize institutional arguments about public accountability, we include public and nonprofit sectors. To operationalize the argument that public and nonprofit unionization drives of the 1970s popularized labor relations offices there, we look at industry by period effects. With unionization controlled, public and nonprofit status should have positive effects in 1971–80. The omitted sectors are printing, transportation, trade, banking, and hospitals.

The Time Trend

Figures 1 and 2 show striking increases in the popularity of all offices but labor relations. Between 1955 and 1985, the prevalence of personnel, benefits, and health and safety offices more than doubled, with the largest gains coming after new federal employment rights were created in the early 1970s. The first EEO/AA offices and officers appeared in 1968, and,

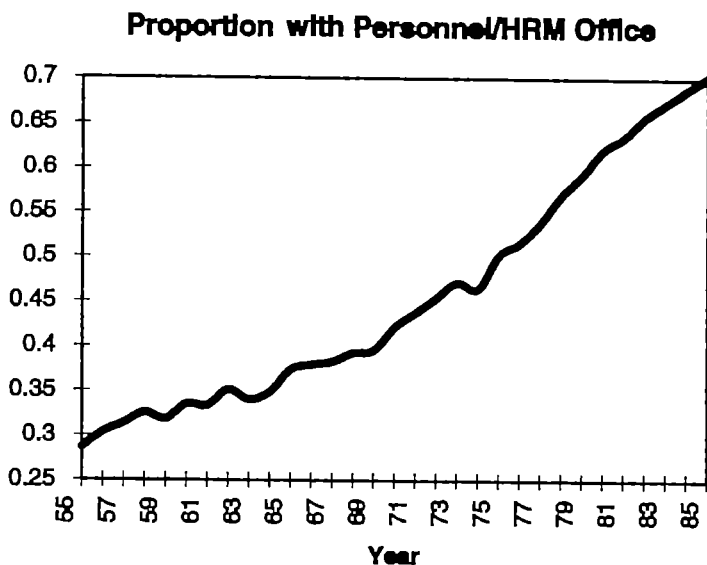


FIG. 1.—Proportion of employers with personnel/human resources office

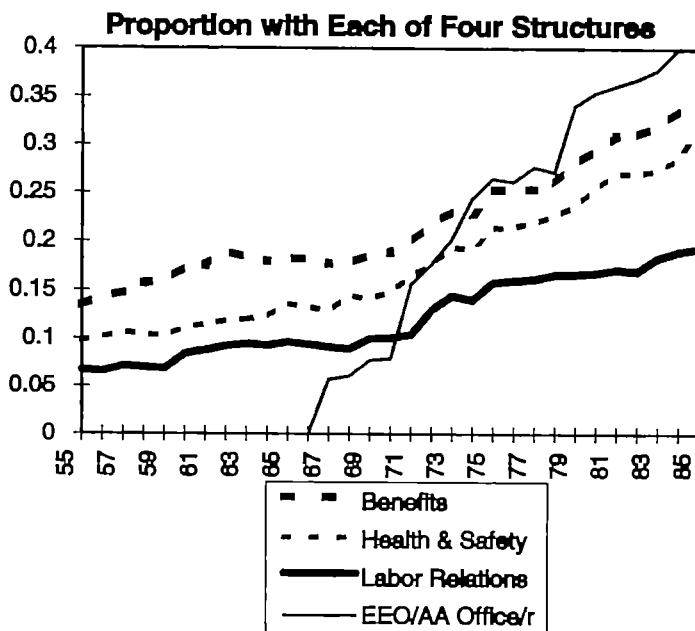


FIG. 2.—Proportion of employers with each of the four structures

by 1985, fully 40% of employers had them. While we can only make tentative inferences from the raw plots, they appear to support our argument about the rights revolution. For personnel, benefits, health and safety, and EEO/AA units, the slopes change notably at the end of the 1960s. For each of these, the slope does not decline after 1980, when regulatory effort declined.

Estimation

Table 2 presents piecewise exponential models, estimated with maximum likelihood techniques using Nancy Tuma's (1979) RATE program. Over the 30-year time frame we cover, the 279 organizations yield 6,701 annual spells of observation (organization-years). Organizations are excluded from the risk set before their birth and after they have adopted a practice, thus, those that had the modeled unit in 1955 are left-censored. The number of at-risk spells ranges from 3,506 to 5,779, and the number of transitions ranges from 41 to 146.

Piecewise exponential models allow us to present a vector of variables hypothesized to have stable effects over time, constants for each time period to pick up changes in the legal environment, and interactions between time periods and other variables. To evaluate period effects, we constrained the coefficient for the constant in period 1 to equal zero.

FINDINGS

In table 2 we present one equation for each of the outcomes of interest. For comparability, we present parallel models for each of the outcomes, varying only the relevant time periods. For simplicity, we include variables that showed *consistent* effects across periods in the first, cross-period, vector. We include variables that showed *inconsistent* effects across periods for *any* outcome in the time-dependent vectors.

Legal Landmarks of the 1970s

The legal changes of the early 1970s had marked effects on employers. For personnel, benefits, and EEO/AA units, the second period, representing a legal landmark, shows a significant positive coefficient. Additional support for the role of regulation is found in our analysis of labor relations offices. Was the creation of new offices simply part of a secular trend toward differentiation? There were no regulatory landmarks that might have spawned labor relations offices, and as we predicted, there is no evidence of a general rise.

Regulatory Effort versus the Reconstruction of Departments as Efficient

The Reagan era shows a significant positive effect. Comparisons of constant terms for periods 2 and 3 show that the diffusion of personnel, benefits, health and safety, and EEO/AA offices did not slow after 1980. For three of the offices, adoption rose in the 1970s and continued at about the same rate (net of the effects of covariates) after 1980. Adoption of health and safety offices rose only after 1980. These findings show that the weakening of regulation during the Reagan years did not slow adoption, as the regulatory effort hypothesis predicts. They support our argument that these offices had been successfully retheorized in terms of efficiency. We find not simply that existing offices survived the first Reagan term but that offices continued to be established apace.

Efficiency Rationales and the Declining Importance of Covariates

Further evidence that these offices had been retheorized in efficiency terms comes from changes in the effects of covariates. Among the four structures stimulated by the rights revolution, there is a decline over time in the importance of functional needs and susceptibility to the law. In the first period, before legal landmarks, a total of 15 coefficients for covariates are statistically significant. In the second period, a total of six are significant. In the third period, only one is significant. As these departments were reconstructed in terms of their generic efficiency, organizations of all sorts adopted them.

Middle Managers as Purveyors

We argued that professional groups promoted new offices. Having a legal office increases the chance of creating each structure in the first period and two of them in the second period. Having a labor/employment attorney on retainer shows generally positive effects. Having a labor relations office increases the chance of adopting benefits and health and safety offices, and this effect does not vary over time. Previous studies suggested that personnel managers oppose the creation of competing departments, yet above, we presented evidence that their associations promoted new positions to handle EEO, health and safety, and benefits. Accordingly, in table 2, we show that employers that belong to personnel associations are more likely to create benefits and EEO/AA units but that employers with personnel departments are no more likely to create EEO/AA and safety units and are *less* likely to create benefits units.

TABLE 2
ESTIMATES OF THE CREATION OF FIVE STRUCTURES AMONG 279 AMERICAN ORGANIZATIONS, 1955-86

Outcome Variable	Personnel/HRM Office	Benefits Office	Health and Safety Office	EEO/AA Office or Officer	Labor Relations Office
Cross-period.					
Constant	-6.22*** (57)	-6.31*** (61)	-6.71*** (63)	-6.84*** (65)	-7.22*** (87)
Log employment	.26*** (07)	.26*** (09)	.26*** (08)	.24*** (07)	.25* (09)
Union	-.45 (30)	-.29 (38)	-.39 (35)	.15 (26)	.89* (37)
Personnel/HRM office		-.63* (31)	-.31 (32)	.04 (15)	-.58* (24)
Labor relations office	85 (48)	1.36*** (48)	1.88*** (38)	-.16 (17)	
Personnel association	.50 (32)	.95*** (32)	.53 (29)	.54* (23)	-.31 (30)
Age of organization	-.6e-3** (2e-2)	-.2e-2*** (4e-2)	-.1e-2*** (4e-2)	-.2e-2 (2e-2)	-.2e-2 (3e-2)
Period 1:					
	1955-69	1955-73	1955-69	1961-71	1955-69
Constant	0	0	0	0	0
Legal office	2.10*** (50)	1.45*** (46)	1.85*** (46)	2.02*** (51)	1.05 (76)
Labor attorney on retainer	89 (48)	.15 (47)	.58 (42)	1.01* (50)	61 (74)
Sector:					
Electrical and machinery	1.64** (60)	.75 (68)	-.69 (108)	-.06 (69)	-8.49 (79)
Chemicals	1.30 (109)	1.99*** (71)	1.39* (60)	1.61*** (54)	1.19 (118)
Local government	1.05 (57)	1.50* (63)	.68 (58)	-.1e+1 (7e+1)	-20 (119)
State or federal	1.89*** (60)	1.20* (62)	-.03 (67)	-.1e+1 (8e+1)	1.28 (86)
Nonprofit	2.16*** (58)	1.60* (64)	1.67*** (62)	-.1e+1 (1e+2)	49 (120)

Period 2:		1970-80	1974-80	1970-80	1972-80	1970-80				
Constant	2.09***	(.51)	1.37*	(.58)	.53	(.70)	1.91**	(.59)	-.33	(.98)
Legal office	.80*	(.41)	.36	(.55)	.31	(.58)	1.01*	(.31)	91	(.57)
Labor attorney on retainer	.68*	(.29)	.84	(.46)	.88	(.48)	.42	(.28)	88	(.47)
Sector:										
Electrical and machinery	.22	(.33)	- 74	(.77)	.69	(.72)	- 27	(.44)	1.28	(.92)
Chemicals	-.26	(.61)	.82	(.68)	1.22	(.90)	-1.13	(1.04)	-.13	(.08)
Local government	.54	(.36)	.01	(.62)	.72	(.79)	.30	(.37)	1.79*	(.86)
State or federal	.21	(.45)	-1.30	(1.07)	1.35*	(.68)	1.31***	(.33)	2.86***	(.83)
Nonprofit	.78*	(.36)	.51	(.80)	1.27	(.73)	- 70	(.51)	1.85*	(.90)
Period 3:										
	1981-85	1981-85	1981-85	1981-85	1981-85	1981-85				
Constant	2.06***	(.57)	1.36*	(.65)	1.66***	(.63)	1.52*	(.73)	.16	(1.10)
Legal office	.22	(.92)	-1.45	(.84)	-1.09	(.82)	1.14	(.84)	83	(.87)
Labor attorney on retainer	.78	(.55)	1.23*	(.54)	.63	(.52)	-.06	(.64)	1.25	(.89)
Sector:										
Electrical and machinery	.22	(.55)	-1.38	(1.07)	-.81	(.80)	.29	(.74)	-1e+1	(1e+1)
Chemicals	.80	(.68)	-.10	(2e+7)	.09	(1.09)	-.08	(1.13)	.60	(1.25)
Local government	-.71	(1.07)	.56	(.68)	.69	(.71)	-.78	(1.15)	.54	(1.09)
State or federal	.51	(.63)	.83	(.63)	.71	(.62)	1.12	(.79)	1.46	(1.05)
Nonprofit	-.07	(1.07)	-.11	(2e+7)	-.11	(2e+7)	-.97	(1.13)	-1e+1	(1e+3)
Chi-square	154.18		100.84		114.66		226.31		78.59	
Number at risk	235		260		256		279		267	
Left-censored cases	44		19		23		0		12	
Number created	146		74		72		111		41	
At-risk spells (N)	3,506		5,111		5,388		4,728		5,779	

NOTE.—Nos in parentheses are SEs.

* $P < .05$

** $P < .01$.

*** $P < .001$.

Age and Public Accountability

Other institutional hypotheses fare well. Older organizations are less likely to create personnel/HRM, benefits, and health and safety offices, as the Stinchcombe/Selznick hypothesis suggests. Public and nonprofit employers are more likely to adopt these same three offices before legal shifts, which suggests that they are most likely to anticipate new legal norms. After legal landmarks, public and nonprofit employers remain more likely to adopt personnel, health and safety, and EEO/AA units.

Scale Economies, Unionism, Transaction Costs, and Segmentation

The other theories show mixed results. Size, measured as log employment, shows robust effects on all five outcomes, lending support to functional arguments. The union drive of the 1970s increases labor relations offices among public and nonprofit employers, even with current union status in the equation. This can be seen in the period 2 effects of local government, state and federal government, and nonprofits on labor relations offices. The strength of these sector effects, even with establishment unionization controlled, suggests that it was not the signing of a union contract that produced new offices but wider union gains. It appears that the growth of labor relations in the 1970s, evident in figure 2, is a consequence of union gains among public and nonprofit employers alone, because the constant term for period 2 shows no effect.

Transaction costs and labor segmentation theories gain some support here. In the first period, before legal landmarks, firms in capital intensive industries that depend on firm-specific skills—electrical manufacturing, machinery, and chemicals—are more likely to install benefits, health and safety, and EEO/AA structures. This effect declines over time, however, as predicted.

CONCLUSION

The federal employment rights revolution of the early 1970s was designed to expand rights to equal protection, life and liberty, and social citizenship in the workplace. We have argued that the ambiguity of compliance standards, expanding scope of the law, and fragmentation of administration led organizations to establish specialty departments to signal a commitment to compliance, to import and invent compliance strategies, and to handle federal regulators. Organizations created new offices not because the law dictated that they do so but because the law did not tell them *what* to do. Management journals defined specialized offices, with expert staffs, as the best protection against costly lawsuits. Our analyses show

that shortly after each legal landmark, employers began to create new offices to handle compliance.

The strength of the weak U.S. state is evident in what happened next. Justifications for these offices drifted away from compliance and toward pure efficiency. By 1980, managers were advocating each kind of office in the rational language offered by the new human resources management paradigm. We argue that this occurred in part because constitutional protections against the rise of federal tyranny signal that government domination of private enterprise is inefficient and illicit. It occurred in part because the law gave employers a particularly active role in devising compliance mechanisms, and hence, middle managers came to see those mechanisms as their own. And it occurred in part because ambiguous federal mandates are a thin reed on which to build elaborate new personnel programs.

It is not that managers deny that the law affects them. It is, rather, that they develop efficiency rationales for the offices they establish in response to the law. By the early 1980s, the new human resources management movement was championing diversity as the key to expanding markets and improving innovation, safety and health programs as the key to winning employee commitment and renovating antiquated technologies, and benefits programs as a means to reducing alienation and improving worker attitudes.

Our data show dramatic effects of this retheorization. When Reagan cut enforcement in 1981, the diffusion of these offices did not falter. As further evidence that departments handling benefits, health and safety, and equal employment opportunity had become disassociated from policy, over time, the policy-related covariates that predict adoption declined in importance. Whereas at first, employers with legal offices, those depending on the primary labor market, and those in public and nonprofit sectors were establishing these offices, by the 1980s, all sorts of employers were doing so.

Does this mean that the employment rights revolution has been an unmitigated success? While these offices continued to diffuse even during the Reagan years, there is evidence that cuts in enforcement led employers to circumvent the law. Employers raided pension funds in response to weakened enforcement of ERISA (Achenbaum 1986, pp. 155–56) and cut affirmative action programs in response to weakened enforcement by the OFCCP (see Bureau of National Affairs 1976, 1986). Moreover, due to the nonuniversal coverage and spotty enforcement of these laws, many workplaces continue to have few protections against discrimination, inadequate safety measures, and no pension or health coverage.

We found managers identifying market forces as the source of structures they had adopted to comply with the law, thereby reinforcing the

authority of the market and sapping the authority of the state. This pattern helps to explain the federal government's combination of administrative weakness and normative strength, which has puzzled students of the state (e.g., Evans et al. 1985). We contend that the pattern is found in business regulation more widely, notably in antitrust law. After antitrust law forced firms to adopt new business practices, industry invented efficiency explanations for those practices. As James Q. Wilson (1980) argues, the Sherman Antitrust Act of 1890, adopted to protect the rights of small business by undermining cartels, spawned price competition. Business soon replaced the cooperative theory of the modern economy, which defined the cartels of the 1880s as efficient, with a natural selection theory, which defined new price competition as efficient (Dobbin 1994). As Neil Fligstein (1990) argues, the Celler-Kefauver Act of 1950, adopted to undermine the power of vertically integrated firms, caused firms to expand into unrelated fields. Business soon replaced the coordination and control theory of the firm, which defined vertical links as efficient, with portfolio theory, which defined conglomeration as efficient.

This pattern, in which Americans develop collective amnesia about the state's role in shaping private enterprises, may explain the long absence of a theory of the state in organizational sociology. Americans subscribe to the theory that firms operate in a Hobbesian economic state of nature, in which behavior depends very much on managerial initiative and markets and very little on political initiative and law. We argue that this theory of the firm is a consequence of the administrative weakness of the federal state, which in the process of constructing public authority as illicit and policy as inconsequential, constructs market authority as legitimate and managers as prepotent.

REFERENCES

- Achenbaum, W. Andrew. 1986. *Social Security: Visions and Revisions*. Cambridge: Cambridge University Press.
- Akey, Denise S., ed. 1984. *Encyclopedia of Associations*. Detroit, Mich.: Gale Research.
- American Hospital Association. 1983. *Hospitals*. Chicago: American Hospital Association.
- Anderson, C. Richard. 1975. *OSHA and Accident Control through Training*. New York: Industrial Press.
- Bardach, Eugene, and Robert A. Kagan. 1982. *Going by the Book: The Problem of Regulatory Unreasonableness*. Philadelphia: Temple University Press.
- Baron, James N., Frank R. Dobbin, and P. Devereaux Jennings. 1986. "War and Peace: The Evolution of Modern Personnel Administration in U.S. Industry." *American Journal of Sociology* 92:350-83.
- Bixby, Ann Kallman. 1986. "Social Welfare Expenditures, 1963-83." *Social Security Bulletin* 49:12-24.

- Blau, Peter 1970. "A Formal Theory of Differentiation in Organizations." *American Sociological Review* 35:201-18.
- Blau, Peter M., Cecilia McHugh Falbe, William McKinley, and Phelps K. Tracy. 1976. "Technology and Organization in Manufacturing." *Administrative Science Quarterly* 21:20-40.
- Blau, Peter M., and Richard A. Schoenherr. 1971. *The Structure of Organizations*. New York: Basic.
- Boyle, M. Barbara. 1973. "Equal Opportunity for Women Is Smart Business." *Harvard Business Review* 51:85-95.
- Bradshaw, David. 1987. "Sexual Harassment: Confronting the Troublesome Issues." *Harvard Business Review* 32:50-52.
- Brandes, Stuart D. 1976. *American Welfare Capitalism, 1880-1940*. Chicago: University of Chicago Press.
- Brickman, Ronald, and Sheila Jasanoff. 1980. "Concepts of Risk and Safety in Toxic Substances Regulation: A Comparison of France and the U.S." *Policy Studies Journal* 9:394-409.
- Burack, Elmer H., and Robert D. Smith. 1977. *Personnel Management: A Human Resource Systems Approach*. New York: West Publishing.
- Bureau of National Affairs. 1976. *Equal Employment Opportunity: Programs and Results*. PPF Survey no. 112. Washington, D.C.: Bureau of National Affairs.
- . 1986. *EEO Policies and Programs*. PPF Survey no. 141. Washington, D.C.: Bureau of National Affairs.
- Burstein, Paul. 1985. *Discrimination, Jobs, and Politics: The Struggle for Equal Employment Opportunity in the United States since the New Deal*. Chicago: University of Chicago Press.
- Burstein, Paul, and Kathleen Monaghan. 1986. "Equal Employment Opportunity and the Mobilization of Law." *Law and Society Review* 16:355-88.
- Chayes, Antonia Handler. 1974. "Make Your Equal Opportunity Program Court-Proof." *Harvard Business Review* 52:81-89.
- Coleman, Barbara J. 1985. *Primer on Employee Retirement Income Security Act*. Washington, D.C.: Bureau of National Affairs.
- Cox, Taylor H., and Stacey Blake. 1991. "Managing Cultural Diversity: Implications for Organizational Competitiveness." *Academy of Management Journal* 5.45-56.
- Cummins, Richard J. 1986. "The General Principles of Law, Separation of Powers, and Theories of Judicial Decision in France." *International and Comparative Law Quarterly* 35.592-628.
- Davey, Patrick J. 1978. "Current Directions in Pension Fund Management." *Conference Board Information Bulletin*, no. 39. New York: Conference Board.
- DiMaggio, Paul J., and Walter W. Powell. 1983. "The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields." *American Sociological Review* 35:147-60.
- Dobbin, Frank. 1994. *Forging Industrial Policy: The United States, Britain, and France in the Railway Age*. New York: Cambridge University Press.
- Dobbin, Frank, Lauren Edelman, John W. Meyer, W. Richard Scott, and Ann Swidler. 1988. "The Expansion of Due Process in Organizations." Pp. 71-100 in *Institutional Patterns and Organizations: Culture and Environment*. Cambridge, Mass.: Ballinger.
- Dobbin, Frank, John Sutton, John Meyer, and W. R. Scott. 1993. "Equal Opportunity Law and the Construction of Internal Labor Markets." *American Journal of Sociology* 99:396-427.
- Doeringer, Peter B., and Michael J. Piore. 1971. *Internal Labor Markets and Manpower Analysis*. Lexington, Mass.: Heath.
- Duhamel, Olivier. 1987. "Remarques sur la notion du régime semi-présidentiel." Pp. 581-90 in *Droits, institutions et systèmes politiques. Mélanges en hommage à Mau-*

- rice Duverger*, edited by Dominique Colas and Claude Emeri. Paris: Presses Universitaires de France.
- DuRivage, Virginia. 1985. "The OFCCP under the Reagan Administration. Affirmative Action in Retreat." *Labor Law Journal* 36:360-68.
- Edelman, Lauren. 1990. "Legal Environments and Organizational Governance: The Expansion of Due Process in the American Workplace." *American Journal of Sociology* 95:1401-40.
- . 1992. "Legal Ambiguity and Symbolic Structures: Organizational Mediation of Civil Rights Law." *American Journal of Sociology* 97:1531-76.
- Edelman, Lauren B., Steven E. Abraham, and Howard S. Erlanger. 1992. "Professional Construction of the Legal Environment: The Inflated Threat of Wrongful Discharge Doctrine." *Law and Society Review* 26:47-83.
- Evans, Peter, Dietrich Rueschemeyer, and Theda Skocpol, eds. 1985. *Bringing the State Back In*. London: Cambridge University Press.
- Ewing, David W. 1983. "Your Right to Fire." *Harvard Business Review* 61:32-52.
- Farley, Jennie. 1979. *Affirmative Action and the Woman Worker*. New York: Amacom.
- Feld, John. 1984. *Affirmative Action: A Fresh Look at the Record Twenty-two Years after the Beginning*. Washington, D.C.: Center for National Policy Review.
- Fligge, Harry E., Jr. 1981. "Defusing the Pension Liability Bomb." *Harvard Business Review* 69:157-87.
- Fligstein, Neil. 1990. *The Transformation of Corporate Control*. Cambridge, Mass.: Harvard University Press.
- Foulkes, Fred K. 1980. *Personnel Policies in Large Nonunion Companies*. Englewood Cliffs, N.J.: Prentice-Hall.
- Foulkes, Fred K., and Henry M. Morgan. 1977. "Organizing and Staffing the Personnel Function." *Harvard Business Review* 55:142-77.
- Froehlich, Herbert, and Dennis Hawyer. 1974. "Compliance Spinoff: Better Personnel Systems." *Personnel* 51:62-69.
- Gavini, Christine. 1997. "Vers un droit interne d'entreprise?" *Sociologie du Travail* 39:149-69.
- Giblin, Edward J., and Oscar Ornati. 1974. "A Total Approach to EEO Compliance." *Personnel* 51:32-43.
- Gill, Kathleen D. 1985. *ERISA: The Law and the Code*. Washington, D.C.: Bureau of National Affairs.
- Gordon, David M., Richard Edwards, and Michael Reich. 1982. *Segmented Work, Divided Workers*. London: Cambridge University Press.
- Greene, Robert J., and Russell G. Roberts. 1983. "Strategic Integration of Compensation and Benefits." *Personnel Administrator* 28:82.
- Hakala, Donald R., and Kenneth M. Huggins. 1976. "ERISA: Impact on Business Financial Management." *Michigan Business Review* 28:19-23.
- Hamilton, Gary G., and John R. Sutton. 1989. "The Problem of Control in the Weak State: Domination in the United States, 1880-1920." *Theory and Society* 18:1-46.
- Hartnett, John. 1996. *OSHA in the Real World: How to Maintain Workplace Safety while Keeping Your Competitive Edge*. Santa Monica, Calif.: Merritt.
- Harvard Law Review*. 1989. "Rethinking Weber: The Business Response to Affirmative Action." *Harvard Law Review* 102:658-71.
- Jackall, Robert. 1983. "Moral Mazes: Bureaucracy and Managerial Work." *Harvard Business Review* 61:118-58.
- Jacoby, Sanford. 1985. *Employing Bureaucracy: Managers Unions, and the Transformation of Work in American Industry, 1900-1945*. New York: Columbia University Press.
- Johns, Horace, and H. Ronald Moser. 1989. "Where Has EEO Taken Personnel Policies?" *Personnel* 66:63-66.

- Kalleberg, Arne L., Peter V. Marsden, Howard E. Aldrich, and James W. Cassell. 1990. "Comparing Organizational Sampling Frames." *Administrative Science Quarterly* 35:658-88.
- Klein, Alfred. 1986. "Why You Can't Afford to Ignore ERISA." *Personnel Journal* 65:72-83.
- Klein, Michael F., Jr., and Robert H. Moses. 1974. "Pension Plans: What's at the End of the Rainbow." *Price Waterhouse and Company Review* 19:10-17.
- Kochan, Thomas A., Harry C. Katz, and Robert B. McKersie. 1986. *The Transformation of American Industrial Relations*. New York: Basic Books.
- Levy, Michel. 1974. "La Politique sociale française et les personnes handicapées, acquis et perspectives." *Analyse et Prevision* 18 243-316.
- Lincoln, James R., and Arne L. Kalleberg. 1985. "Work Organization and Workforce Commitment: A Study of Plants and Employees in the U.S. and Japan." *American Sociological Review* 50:738-60.
- Logue, Dennis E. 1979. *Legislative Influence on Corporate Pension Plans*. Washington, D.C.: American Enterprise Institute.
- Lovecy, Jill. 1992. "Comparative Politics and the Fifth French Republic. La fin de l'exception française." *European Journal of Political Research* 21:385-408.
- Malkiel, Burton. 1979. "Productivity: The Problem behind the Headlines." *Harvard Business Review* 57:81-111.
- Marino, Kenneth. 1980. "Conducting an Internal Compliance Review of Affirmative Action." *Personnel* 57:24-34.
- McCaffery, Robert M. 1986. *Employee Benefit Programs: A Total Compensation Perspective*. Boston: PWS-Kent Publishing.
- Mendelhoff, John. 1979. *Regulating Safety*. Cambridge, Mass.: MIT Press.
- Meyer, John W. 1983. "Centralization of Funding and Control in Educational Governance." Pp. 179-98 in *Organisational Environments: Ritual and Rationality*, edited by John W. Meyer and W. Richard Scott. Beverly Hills, Calif.: Sage.
- Meyer, John W., and Brian Rowan. 1977. "Institutionalized Organizations: Formal Structure as Myth and Ceremony." *American Journal of Sociology* 83:340-63.
- Meyer, John W., and W. Richard Scott. 1992. *Organisational Environments*. Newbury Park, Calif.: Sage Publications.
- Meyer, Marshall. 1979. "Organizational Structure as Signaling." *Pacific Sociological Review* 22:481-500.
- Meyer, Mitchell. 1981. *Profile of Employee Benefits* 1981 edition. New York: Conference Board.
- Meyer, Mitchell, and Harland Fox. 1974. *Profile of Employee Benefits*. New York: Conference Board.
- Mill, John Stuart. 1988 (1974). *A System of Logic*, edited by J. M. Robson. Toronto: University of Toronto Press.
- Moody, Hiram F., Jr., and Edward D. Higgins. 1984. "Selling the 401(k) Plan to Employees." *Harvard Business Review* 62:68-78.
- National Industrial Conference Board (NICB). 1940. *Personnel Activities in American Business*. New York: NICB.
- Northrup, Herbert, Brian Rowan, and Charles Perry. 1978. *The Impact of OSHA*. Labor Relations and Public Policy Series, no. 17. Philadelphia: University of Pennsylvania, Wharton School, Industrial Research Unit.
- Pati, Gopal C., and John I. Adkins, Jr. 1980. "Hire the Handicapped: Compliance Is Good Business." *Harvard Business Review* 58:14-29.
- Petersen, Donald J. 1974. "The Impact of Duke Power on Testing." *Personnel* 51:31-37.
- Posner, Richard A. 1997. "The Rise and Fall of Administrative Law." *Chicago-Kent Law Review* 72:953-63.

- Rees, Joseph 1988. *Reforming the Workplace: A Study of Self-Regulation in Occupational Safety*. Philadelphia: University of Pennsylvania Press.
- Ritter, Anne, and William H. Wager. 1989. "Special Awards for Special Achievers: Personnel Winners." *Personnel* 66:17-32
- Ropp, Kirland. 1987. "How the Courts Affect Management Policy." *Personnel Administrator*. 32:45-50. 100.
- Scott, W. Richard, and John W. Meyer 1987. "Environmental Linkages and Organizational Complexity. Public and Private Schools." Pp. 128-60 in *Comparing Public and Private Schools*, edited by Henry M. Levin and Tom James. New York. Falmer.
- Selznick, Philip. 1949. *TVA and the Grass Roots*. Berkeley: University of California Press.
- . 1969. *Law, Society and Industrial Justice*. New York: Russell Sage
- Shaeffer, Ruth G. 1973. *Nondiscrimination in Employment: Changing Perspectives, 1963-1972*. New York: Conference Board.
- Skolnik, Alfred M. 1976. "Twenty-five Years of Employee-Benefit Plans." *Social Security Bulletin* 37 (3). 3-22.
- Skowronek, Stephen. 1982. *Building a New American State: The Expansion of National Administrative Capacities, 1877-1920*. New York: Cambridge University Press.
- Skrentny, John D. 1996. *The Ironies of Affirmative Action: Politics, Culture, and Justice in America*. Chicago: University of Chicago Press.
- Spann, Jeri. 1990. "Dealing Effectively with Sexual Harassment. Some Practical Lessons from One City's Experience, Madison, Wisconsin." *Public Personnel Management* 19:53-82.
- Stinchcombe, Arthur. 1965. "Social Structure and Organizations." Pp. 142-93 in *Handbook of Organizations*, edited by James March. Chicago: Rand McNally
- Strang, David, and John W. Meyer 1994. "Institutional Conditions for Diffusion." Pp. 100-12 in *Institutional Environments and Organization: Structural Complexity and Individualism*. Edited by W. Richard Scott and John W. Meyer. Thousand Oaks, Calif.: Sage.
- Stryker, Robin. 1996. "Law, Sociology, and Public Policy Issues in Equal Employment Opportunity." Paper presented at the annual meetings of the American Sociological Association, New York.
- Sunstein, Cass R. 1996. "The APA at Fifty: Health-Health Tradeoffs." *University of Chicago Law Review* 63:1533-71
- Sutton, John, and Frank Dobbin 1996. "The Two Faces of Governance: Responses to Legal Uncertainty in American Firms, 1955-1985." *American Sociological Review* 61:794-811.
- Sutton, John, Frank Dobbin, John W. Meyer, and W. Richard Scott. 1994. "The Legalization of the Workplace." *American Journal of Sociology* 99:944-71
- Tepper, Irwin. 1977. "Risk vs." *Harvard Business Review* 55:100-115
- Tolbert, Pamela S., and Lynne G. Zucker 1983. "Institutional Sources of Change in the Formal Structure of Organizations. The Diffusion of Civil Service Reform, 1880-1935." *Administrative Science Quarterly* 28:22-39.
- Tuma, Nancy B. 1979. *Invoking RATE*. Palo Alto, Calif.: DMA Press.
- Vernon-Geerstenfeld, Susan, and Edmund Burke 1985. "Affirmative Action in Nine Large Companies: A Field Study." *Personnel* 62:54-60.
- Williamson, Oliver. 1975. *Markets and Hierarchies: Analysis and Antitrust Implications*. New York: Free Press.
- Wilson, James Q., ed. 1980. *The Politics of Regulation*. New York: Basic.
- Wooten, James A. 1994. "Regulating the 'Unseen Revolution': Private Pension Policies and Politics, 1960-1987." Working paper. SUNY-Buffalo, School of Law.
- Zippo, Mary. 1980. "Roundup." *Personnel* 57:65-66

High-Risk Collective Action: Defending Human Rights in Chile, Uruguay, and Argentina¹

Mara Loveman

University of California, Los Angeles

Under what conditions will individuals risk their lives to resist repressive states? This question is addressed through comparative analysis of the emergence of human rights organizations under military dictatorships in Chile, Uruguay, and Argentina. While severe state repression is expected to lead to generalized demobilization, these cases reveal that repression may directly stimulate collective action. The potential for sustained collective action in high-risk contexts depends upon the relationship between strategies of repression and the particular configuration of embedded social networks; it is more likely where dense yet diverse interpersonal networks are embedded within broader national and transnational institutional and issue networks.

First, we will kill all the subversives; then we will kill their collaborators; then their sympathizers; then the indifferent and finally, the timid. (attributed to Brig. General Ibérico Manuel Saint-Jean, former governor of Buenos Aires)

INTRODUCTION

From the mid-1960s into the late 1980s, military governments in Latin America systematically and ferociously violated human rights as defined by the international community with the adoption of the Universal Decla-

¹ I would like to thank the following individuals for their comments on earlier versions of this article: Rogers Brubaker, Rebecca Emigh, Elizabeth Lira, David Lopez, Brian Loveman, Michael Mann, Gerardo Munck, Jeffrey Prager, William Roy, Caleb Southworth, Jay Sundu, Ivan Szelenyi, and Maurice Zeitlin. I also wish to thank the participants in the UCLA Comparative Social Analysis Workshop, fall 1997, and the *AJS* reviewers. Revisions were supported in part by the Andrew A. Mellon Foundation. Direct Correspondence to Mara Loveman, Department of Sociology, University of California, 264 Haines Hall, Box 951551, Los Angeles, California 90095.

ration of Human Rights in 1948. Barbaric torture, murders, "disappearances,"² and disregard for civil liberties and rights became commonplace. In response, individuals and small groups of citizens protested against these policies, called for changes in government behavior, and sought to support and assist victims. Human rights organizations (HROs) formed, and a diffuse human rights movement arose in most of Latin America.

In some countries, HROs formed quickly and had significant social support. In others, HROs developed relatively later, were more fragile, and were less influential. In all cases, however, those participating in HROs faced potential retribution, even death, for opposing the incumbent dictatorships.

This article seeks to explain the emergence of sustained collective action in such "high-risk" contexts. Specifically, it attempts to account for the emergence of HROs in Chile, Uruguay, and Argentina in contexts where the potential consequences included arrest, torture, disappearance, or murder of participants, their friends, and their family members.³ In short, why and how did individuals organize resistance in the face of extremely repressive governments? Further, what accounts for the significant differences in the timing of the emergence of HROs in these countries?

Asking these questions draws attention to the paucity of explicit attempts in the social movement literature to explain the emergence of sustained collective action in contexts of life-threatening risk. Most social movement theory stems from research into low-risk forms of mobilization. However, as McAdam (1986) suggests, the mobilization dynamics of high-risk movements are likely to be qualitatively different from those of low-risk movements. This article suggests that explanation of the emergence of HROs in Chile, Uruguay, and Argentina requires a synthetic theoretical approach that focuses attention on links between interpersonal and embedded social and political networks, resource mobilization capacity, identity construction prior to and in the process of participation, and the

² The word "disappearances" refers to kidnappings and murders not acknowledged by governments.

³ The central empirical focus of this article is the human rights organizations (HROs) that emerged in each country. Recognizing that social movement organizations are only one possible form of the *organisation of collective action* in a social movement (Tarrow 1994, p. 136), this account focuses on HROs for two reasons: first, available accounts of the early years of these movements by participants and scholars suggest that they consisted first and foremost of the networks of individuals working under the auspices of HROs. Second, these accounts also indicate that HROs provided the "mobilizing structures" (Tarrow 1994, p. 136) that linked diverse elements of the movement domestically and internationally.

modalities and extent of state repression that shape the political opportunity structure.⁴

THEORETICAL APPROACH

This investigation on human rights organizations in Latin America in the 1970s and 1980s asks two basic and interrelated theoretical questions: (1) When (under what conditions) do high-risk social movements and organizations emerge? (2) Why do people participate in such movements and organizations?

Different levels of abstraction and varying levels of theory yield distinct "explanations" for collective action, of which high-risk social movements and organizations are particular instances.⁵ Theories operating at different levels of abstraction are generally presented in competition with other "schools" or "paradigms," although, often, "competing" theories present more of a contest between "most important variables" than between more profound paradigmatic differences. Rather than insist on the uniform predominance of micro, meso, macro, or "supra" macro variables or on variables emphasizing human agency or structural and conjunctural variables, the richness of social movement theories provides the possibility of improving explanations for social movements by reference to *composite causal contingencies* (Lofland 1996). In short, explanations for collective action involve multiple variables whose influence in particular instances of collective action is complexly and contingently interrelated.

SOCIAL MOVEMENT THEORIES AND HIGH-RISK COLLECTIVE ACTION

Different social movement theories offer varying answers to the two basic questions posed above. The following overview considers how prevailing

⁴ Recognizing that in comparative historical research there is always a tension between idiosyncratic rigor and nomothetic explanation, a central aim of this article is to provide an initial consideration of how well a careful synthesis of selected theories can account both for the emergence of particular Latin American HROs and for differences among them.

⁵ The social movements literature is replete with competing definitions of "social movement" (leading Marwell and Oliver [1984, p. 4] to conclude "the concept 'social movement' is a theoretical nightmare"). This article adopts a working definition based on Tarrow's (1994, p. 11) conceptualization: "collective challenges by people with common purposes and solidarity in sustained interaction with elites, opponents and authorities." Adopting this definition does not imply that I accept Tarrow's theoretical arguments regarding the sociological bases for collective challenges in their entirety.

approaches to these two questions account for the emergence of, and participation in, social movements and organizations *in high-risk contexts*.

Micro-Level Approaches

Approaches based on the premises of methodological individualism focus on the decisions and actions of discrete individuals to explain *why people engage in collective action*. These approaches can be divided into two (not necessarily mutually exclusive) general categories: those that adopt a rational choice framework and those that focus on individual characteristics or psychology to explain motivation.

Focusing attention on individual calculations of the costs and benefits of participation in collective action (Oberschall 1973, 1993; Klandermans 1984), rational choice approaches help explain why individuals participate in some movements but not others, and at some moments but not others, by assessing the changing incentives and disincentives for participation. While accepting Olson's (1965) basic framework, many rational choice theorists have qualified (or relaxed) his assumptions, expanded their analyses from individual to "collective" decision making, and introduced non-material incentives into calculations of cost/benefit ratios (Oliver 1984; Fireman and Gamson 1979; Oberschall 1993). Despite such efforts, both formal and more "relaxed" rational choice models are particularly *unhelpful* for explaining participation in collective action in situations involving high levels of risk or contexts of extreme instability and unpredictability.

Rational choice models work best when the "rules of the game"—the costs and benefits of choosing one action versus another—are clear and predictable (Geddes 1995, p. 87). But many "high-risk" contexts are high risk precisely because the consequences of actions are impossible to predict.⁶ Additionally, it is difficult for rational choice models to avoid tautology when they attempt to incorporate noninstrumental incentives or to explain what Weber ([1992] 1968, p. 25) referred to as "value-rational behavior." In such cases, the "preferences" of actors are often deduced from the very actions they are meant to explain (e.g., Fireman and Gamson 1979, p. 24).

In part, this is because the language and grammar of rational choice tends to obstruct rather than facilitate understanding of human action in contexts where nonmaterial incentives, or "meanings," are *central* to individual decisions. It is analytically useful to distinguish "costs" ("expen-

⁶ The military regimes in Southern Cone countries pursued (to varying degrees) deliberate strategies of seemingly arbitrary and disproportionately cruel forms of repression, creating fear, confusion, and terror with the intention of paralyzing any opposition.

ditures of time, money, and energy required of a person engaged in a particular form of activism") and "risks" ("anticipated dangers of engaging in a particular type of activity") (McAdam 1986, p. 67). However, in the rational choice framework, psychosocial processes such as *fear* are not formally theorized. Hence, risks simply weigh in as potential "costs" in individual calculations. And if risk or cost is calculated as a high probability of "death," while benefit is calculated at a minimal probability of "maintenance of honor" or "respect for human rights," how is this "ratio" to be assessed in the grammar of rational calculation in order to predict the outcome? If the likely result of action is death, rational choice models would predict inaction, unless they determine *ex post facto*, with reference to the individual's behavior, that the first order preference is a certain "value" that requires such a sacrifice. This, of course, is tautological.

Combinations of material and nonmaterial incentives are usually involved in individual decision making, and the reasons for individual participation in social movements may change over time. Different types of social movements, different moments in a movement's development, and different stages and levels of individual participation may be characterized by different combinations of material, solidary, and purposive motivations (Hirsch 1986; Snow et al. 1986). It is possible that at some moments of participation in high-risk social movements, material self-interest may become increasingly important for some participants (i.e., for individuals who become "movement entrepreneurs"); in such instances, a rational choice approach may have considerable explanatory potential. In general, however, it is likely that solidary and purposive motivations usually outweigh material incentives in decisions to participate in high-risk, and especially life-threatening, collective action.⁷

In addition to rational choice approaches, numerous other theories operating at the "micro" level of analysis also focus on individual motivations for participation but offer alternative motivational rationale. Some theorists emphasize how "prosocial" commitments motivate individuals to act collectively to achieve social goals (Martín-Baró 1983). Many "new social movement" (NSM) theorists argue that the expression of "new" or previously unarticulated "identities" forms the basis for collective action. Such approaches may be helpful for explaining participation in contexts of high risk if they avoid essentializing "identity" and instead are sensitive to social processes of identity construction leading up to and in the process of

⁷ See, e.g., Calhoun (1991). Additionally, recent studies emphasizing the role of social networks and processes of collective identity construction suggest that material incentives may be less important determinants of participation than was once thought, even in less risky contexts (Gould 1995; McAdam 1986; Snow and Benford 1992; Snow et al. 1986). Of course, some high-risk collective action is motivated primarily by material incentives

struggle and how socially constructed "identities" may be threatened by social change.

This contribution of NSM theorists may be linked to approaches that explore how "relative deprivation" or "status strain" may induce participation in social movements (Gurr 1970; Lipset and Raab 1970). The insights of these contributions can be applied to contexts of high risk by considering how sudden social change affects the self-understandings of particular individuals or groups in specific ways. Sudden threats to lifestyle or fundamental values may make life-risking behavior appear like the only viable option—as "self-saving" rather than "self-sacrificing" (Calhoun 1991, p. 69; Weber [1922] 1968, p. 25). Under such conditions, taking high risks may seem "the only choice."

Constructivist Approaches

A recurrent criticism of motivational accounts is that they fail to make problematic the social processes, or "micromechanisms," through which shared grievances or "latent" motivations are translated into collective action. One line of this criticism argues that explanations of participation in collective action need to consider how interactive processes of *interpretation* of grievances influence the propensity to participate (Snow et al. 1986, p. 465). Other theorists emphasize the importance of "identity" to social movements. However, theorists disagree over the extent to which identity *explains*, or is *explained by*, participation in a social movement.

Rather than accepting preexisting identities or feelings of solidarity as causal explanations for participation in social movements, some theorists have emphasized how individual and collective interests, identities, and solidarity are conceived, constructed, maintained, and reproduced in the process of struggle itself (Gould 1995; Melucci 1988; Calhoun 1991; Hirsch 1986; Snow et al. 1986). As Calhoun (1991, p. 52) explains: "The issue of identity is not adequately dealt with in terms of *legitimation*, *expression*, or other terms that imply that it exists prior to and is the basis of a struggle. Identity is, in many cases, forged in and out of struggle, including participation in social movements."

This insight is of particular importance for understanding participation in high-risk social movements, where the nature of the risks taken and sacrifices made are likely to heighten processes of self and collective (re-)definition in ways that have important consequences for continued and intensified participation. While, to the detached observer, the decision to risk life for a cause may appear stupid, pathological, or at best irrational, Calhoun (1991, p. 51) suggests instead that risks may be borne "because participation in a course of action has over time committed one to an identity that would be irretrievably violated by pulling back from risk."

Social Networks

Micromobilization processes are affected by the types of social networks within which potential participants are embedded. This is especially the case for recruitment to potentially risky forms of collective action (McAdam 1986; della Porta 1988; Morris 1984; Laitin 1995). Specifically, as Gould (1995, pp. 203–4) suggests and as McAdam (1986) demonstrates in the case of Freedom Summer volunteers, personal ties such as kinship or close friendship are particularly important for recruitment to high-risk activism. This is supported by Wickham-Crowley's (1989) finding that kinship or patron-client relationships were important bases of recruitment for guerrilla movements in Venezuela, Colombia, and Cuba. Personal ties are particularly important for sustaining contentious collective action in extremely repressive contexts because they provide a foundation for constructing the types of dense and insulated social networks required for effective resistance to state attempts at infiltration. The lack of anonymity also heightens individual accountability and can thus help to prevent defection (Laitin 1995). At the same time, it is within such insulated networks built on ties of kinship and friendship that collective and individual identities are constructed that make life-risking action appear as "self-saving" (Calhoun 1991, p. 69).

In addition to the importance of personal social networks, connections to certain types of organizations or institutions are also consequential for recruitment to high-risk collective action. Local social networks are generally embedded within broader regional, national, and sometimes international institutional networks. Theorists have shown that political parties can be an important basis for recruitment to guerrilla movements (Wickham-Crowley 1989) or terrorist groups (della Porta 1988), both forms of high-risk activism. In certain contexts, linkages to religious institutions, at the local, national, or international level, may also facilitate involvement in high-risk social movements (Wickham-Crowley 1989; Morris 1984; Orellana and Hutchison 1991; Lowden 1996; Osa 1989). The way in which face-to-face social networks are embedded within, and structured by, broader organizational networks influences the likelihood of participation in risky collective action.

Resource Mobilization

Connections to certain types of institutions may also help to explain *when (under what conditions) high-risk social movements and organizations emerge*. As numerous theorists in the "resource mobilization" tradition have convincingly argued, for grievances or beliefs in a cause to be translated into collective action requires availability and access to organizational and other resources (Zald and McCarthy 1979; Tilly 1978; Ob-

erschall 1973). Much like "low-risk" social activists, high-risk activists must have access to basic resources in order to begin and sustain mobilization and act collectively to achieve goals. Institutional and personal linkages to churches (Morris 1984), political parties, labor unions, universities, professional organizations, and national and international nongovernmental organizations (NGOs), among others, may provide potential activists in high-risk contexts with resources such as funding, information, and access to physical and symbolic space, which make sustained collective action possible. These sorts of connections frequently provide the "missing links" between feelings of deprivation and injustice, the "interpretation" of these feelings through social networks and processes of identity construction, and contentious collective action.

Frequently, critical resources are provided by "outsiders"—individuals or organizations other than the aggrieved population (Zald and Mayer 1979; Wickham-Crowley 1989). The case studies that follow suggest the acute importance of external financial or material support, information and expertise (services), and access to physical, sociopolitical, and symbolic space for HROs to emerge and be sustained under repressive military dictatorships.

Political Opportunity Structure

Access to resources, existing social networks, and availability of sociopolitical space are not sufficient to account for *when* social movements emerge. To explain the timing of social movement ebbs and flows, theorists have focused on how the political opportunity structure and long-term cycles of protest largely determine the possibility of contentious collective action (Tilly 1978; Tarrow 1994; Diani 1996). Social movements form in response to changes in society and in the political opportunity structure (Tarrow 1994, p. 18). In Tarrow's formulation, political opportunity structure is "external" to potential collective actors, so even resource-poor, "weak or disorganized challengers" can take advantage of changing opportunities. External changes may reduce the costs of collective action, reveal potential allies, and expose the vulnerabilities of authorities (Tarrow 1994, p. 18).

Changes in the political opportunity structure may go a long way toward explaining the timing and tactics of high-risk social movements.⁸ However, it may not always be *improvements* in political opportunity that

⁸ The inherent ambiguity of "changes in the political opportunity structure" may also introduce methodological difficulties. "Opportunity" is highly subjective, and "changes" in opportunity are necessarily relative. Looking back, it is possible to see "opportunities" where movement activists saw only "constraints" (and vice versa) (see Tilly 1978, p. 7).

induce social movement participation. In some cases, state policies and actions may themselves *create* new social movement "constituencies." In the cases examined here, political repression and human rights violations created new constituencies in a particularly dramatic and terrible fashion. Thus, while it may be the case that severe state repression tends to correspond with generalized demobilization (Tilly 1978), excessive abuses by the state may *directly stimulate* the emergence of *certain types* of contentious collective action. A basic assumption of opportunity structure and resource mobilization arguments, that "grievances" remain relatively constant, is violated in cases of sudden and severe state repression. The "early risers" in such contexts may mobilize in response to, *not despite*, severe repression; their actions may then create space for later waves of participants who may indeed be responding to *relative* improvements in the structure of opportunities.

In sum, the insights gained from an overview of theoretical literature on social movements suggest that explanations of participation in high-risk collective action should consider the influence of face-to-face, personal social networks, the way in which immediate social networks are embedded within broader institutional networks, and the forging of increasingly committed self and collective identities in the process of participation. Additionally, in contexts of high risk, social movements and organizations are more likely to emerge when individual and institutional networks facilitate access to critical resources such as information, financial assistance, and physical, sociopolitical, and symbolic space from which to launch contentious challenges. As is true for social movements in general, "external" support enhances the possibility for the emergence of social movements and organizations in high-risk circumstances. Finally, repression (a worsening of the political opportunity structure) may induce certain types of collective action.

EMERGENCE OF HROs IN CHILE, URUGUAY, AND ARGENTINA

Despite significant variations in their political histories and socioeconomic systems, the common experience of military dictatorship characterized by extensive human rights violations invites comparison of the HROs that developed in Chile, Uruguay, and Argentina in the 1970s and 1980s. Of course there were significant differences between the Southern Cone military regimes, as well as in the character of the military governments over time. The particular historical and institutional role of the military in politics differed in each country. The 1973 coups in Chile and Uruguay interrupted long traditions of civilian rule with only a few brief interludes of military involvement. In Argentina, military intervention in politics was a recurrent historical routine. Southern Cone military governments also

varied in their internal structure, from relatively tight, top-down hierarchical control under General Pinochet in Chile, to interforce (army, navy, air force) divisions in Argentina, to hard-line/soft(er)-line factions in Uruguay. Despite such differences among and within these military governments, they shared a common commitment to national security doctrines that articulated the military's historical mission and duty, as ultimate protector of the "patria" from external and internal threats, to eradicate "subversive" elements from society. In all three cases, repression targeted organizations and individuals with linkages or associations with the left, especially labor unions, political parties, and students; however, in no case were government attacks limited to these sectors.⁹ Evidence suggests communication and collaboration among the military intelligence services across nation-state frontiers, including the sharing of novel torture techniques (*Centro de Estudios Legales y Sociales* 1984).

However, the overall repressive strategy of each regime in accomplishing its "mission" varied, from the infamous "dirty war" in Argentina that left as many as 10,000–15,000 people dead or disappeared, to the relatively low number of fatalities but extraordinarily high incarceration and torture rates in Uruguay, where prisons were geared toward psychological torture and systematic destruction of the personality (Weschler 1990). Variations in repressive strategies across countries and over time subjected activists to different sets of risks and constraints in their efforts to create and sustain HROs. Differences existed both in the opportunity structure and the *perception* of risk for human rights activists. Such differences help to account for variations in the participation and timing of HROs in these countries, as the nature and extent of physical and psychological repression confronted by potential participants shaped their perceptions of the risks and opportunities. The timing and extent of legal restrictions on personal liberties and rights also varied across countries.¹⁰ Together with the extent and intensity of repression (see fig. A3 in the appendix), such restrictions provide an indirect measure of the sociopolitical space provided by the "political opportunity structure" at different moments beneath the military regimes in each country.

To compare the political opportunity structure—"dimensions of the political environment that either encourage or discourage people from using collective action" (Tarrow 1994, p. 18)—across countries, this article fo-

⁹ Part of the repression in each country was intended to dismantle labor movements in order to impose neoliberal economic models (see Drake 1996).

¹⁰ Tables listing formal restrictions on political and sociocultural activity in Chile (1973–77), Uruguay (1967–77), and Argentina (1975–76) are available from the author upon request.

cuses on two features of the political environment of particular relevance to potential participants in HROs: (1) the extent and intensity of repression, and (2) formal (legal) restrictions on political and sociocultural space. These two "factors" do not fully operationalize what is meant by the concept "political opportunity," let alone *perceptions* of political opportunity, but together they provide a way to approximate the "space" available in each country and to compare in relative terms the degree of risk confronted by potential participants.

While repressive strategies varied, in all cases HROs emerged in response to sudden and dramatic increases in state abuses of civil and human rights. These were *reactive* social movements, in which human rights activists and constituencies were "created" by the abuses they then sought to end.¹¹ In each country, HROs emerged in contexts characterized by what Corradi, Fagen, and Garretón (1992) refer to as a "culture of fear," where the intentional propagation by a regime of a climate of uncertainty, insecurity, and terror aims to paralyze forms of collective action. Threats of persecution, arrest, torture, disappearance, or assassination of opponents of the regime are meant to create insurmountable obstacles to collective action; they exacerbate existing incentives to free ride.¹² Yet in spite of selective *disincentives* to participate in HROs, such organizations emerged in each of the countries chosen for study.¹³ In this context, how were HROs formed despite both generalized perceptions and actual situations of high risk in Chile, Uruguay, and Argentina? What accounts for cross-country variations?

¹¹ Prior to the 1960s, there was only one organization in the Southern Cone explicitly concerned with human rights: the Liga Argentina por los Derechos del Hombre was created in 1937 in response to political persecution that followed the military coup of Uriburu in 1930 (see Veiga 1985, pp. 15–26; Frühling et al. 1989).

¹² Olson's (1965) formulation of the "free rider" problem, with its emphasis on "proportional participation," does not adequately capture the dynamics of collective action in cases where small numbers may be more of an advantage than a liability (see Tarrow 1994, p. 15).

¹³ The existence of "selective *disincentives*" to participate means that the free rider problem was not overcome through Olson's suggested solution of selective incentives for participants. In some ways, however, the creators and early participants in HROs can be conceptualized as analogous to Olson's "privileged group"—the other possible solution to the free rider problem. While referring to a group of individuals who lived with the constant fear of repression and potentially horrific mental and physical abuse as "privileged" is an obvious misnomer, this group played a role analogous to the "privileged group" in Olson's terminology. For reasons to be explored below, they were willing and able to assume the "costs" and risks of defying the military regime in the hope of providing the "public good" of basic respect for human rights.

CHILE

Rapid Formation of HROs during the Height of Repression

Chile is unique among the cases in the extent to which organized moral opposition (Lowden 1996) to the military regime by individuals not directly affected by human rights abuses began and was sustained during the height of repression.¹⁴ How do the theoretical approaches to social movements reviewed above help account for the emergence of HROs during the worst years of repression (1973–76), and how do they help explain why individuals risked their lives to participate in the earliest HROs in Chile?

The literature on the human rights movement in Chile emphasizes that HROs formed in successive generations, with the first wave based in religious organizations, especially the Catholic Church, the second wave primarily composed of family members of victims, and later waves based in political parties or some combination of these three (Orellana and Hutchison 1991).¹⁵ The creators and participants in the first HROs in Chile were predominantly religious leaders of various denominations and professionals, including academics, politicians, social workers, psychologists, and lawyers. The first organizations were formed in collaboration with international human rights and religious organizations under the protective umbrella of the Catholic Church. Immediately after the coup, the United Nations High Committee for Refugees (UNHCR), the World Council of Churches, through its representative to Chile, and representatives from evangelical churches petitioned the junta for permission to create an organization to evacuate the 10,000–15,000 non-Chilean political refugees residing in Chile and threatened by persecution. The National Committee for Aid to Refugees (CONAR) facilitated the safe exodus of approximately 4,000 refugees by 1974 (Lowden 1996, p. 32). An organization to assist Chileans was created in October of 1973. The Comité de Cooperación para la Paz en Chile (COPACHI), which included Catholic, Jewish, Orthodox, Lutheran, and various evangelical religious leaders, also emerged during the most intense period of repression. Programs under the auspices of COPACHI aimed at providing legal aid and general assistance to vic-

¹⁴ The literature on HROs in the Southern Cone distinguishes between organizations made up primarily of *afectados* or *victimias*, people personally affected by repression through loss of family member, and organizations of *no-afectados*, people indirectly affected by human rights abuses (although there is some overlap in membership).

¹⁵ Some sources refer to groups of family members of *desaparecidos* and victims of human rights abuses as "movements" and reserve the label "organization" for other types of HROs. Since this does not correspond with predominant usage in the social movement literature and would only add to terminological confusion, I have not adopted this usage.

tims of political persecution and their families, assisting refugees, and helping those hardest hit by the regime's austere economic policies through self-help employment projects, day care centers, and children's soup kitchens (Smith 1982, p. 334). In 1975, COPACHI was dissolved under pressure from the military junta and was replaced by the Vicaría de la Solidaridad, which together with La Fundación de Ayuda Social de las Iglesias (FASIC) continued to provide crucial support to those affected by repression.

Subsequent waves of HROs, such as organizations of family members of the disappeared, followed by more explicitly antiregime HROs, emerged in the space opened up by earlier organizations (Vidal 1986, pp. 26–33). The programs and strategies of later waves of HROs varied depending on the political context and particular goals but generally included some combination of public denunciation of human rights abuses, legal and material aid for victims and their family members, popular education, and eventually mass mobilization in opposition to the military regime.¹⁶

Several HROs emerged after 1977, as the dictatorship gradually loosened its control over social life, lending support to Tarrow's thesis that changes in the political opportunity structure that reduce the costs or potential risks of organization largely account for the timing of contentious collective action. Tarrow's formulation helps to account for the emergence of these third-generation Chilean HROs, because it recognizes that *relative* improvements in the opportunity structure may incite collective action. However, the earliest Chilean HROs do not fit neatly in Tarrow's model. What insights does the synthetic theoretical framework outlined above shed on why and how HROs emerged and were sustained during the height of repression in Chile?

The most salient feature of the early HROs in Chile, and the one that distinguishes the Chilean human rights movement from the other cases under consideration, is the prominence of the Catholic Church.¹⁷ But the involvement of the Church per se is not itself an *explanation* for the quick and enduring formation of HROs in Chile compared with the more delayed or relatively weaker responses in other cases. A theoretically in-

¹⁶ Tables listing HROs created in Chile, Uruguay, and Argentina prior to democratic transition in each country are available from the author upon request. Tables include name, date of origin, principal participants, and primary objectives and activities of each organization.

¹⁷ For the purposes of this article, I consider the consequences of the Catholic Church's position vis-à-vis the military regime for HROs without attempting to explain *why* the Church as institution reacted differently to human rights abuses in each country. Though a fascinating question, it is beyond the scope of the present analysis. For Chile, see Smith (1982). For Argentina, see Mignone (1988).

formed assessment of the emergence of and participation in the earliest HROs in Chile suggests specific ways in which the Church and other religious organizations were influential.

First, accounts of the formation of COPACHI reveal the importance of *preexisting transnational and national social and political networks* linking religious leaders from diverse denominations in the creation of the earliest HROs. In the first months of the dictatorship, characterized by massive disappearances and political repression,¹⁸ the church as an institution did not openly criticize the regime. To the contrary, it lent the military government at least tacit support and hence contributed to its legitimacy.¹⁹ Brian Smith (1982) describes the position of the Church during this period of extreme repression as ambiguous at best, with Silva Cardinal Henriquez, the Permanent Committee of the Episcopal Conference, and several bishops expressing confidence in the regime and counting on the "good faith" of the military junta "as Christians" to bring a rapid end to the repression.²⁰

In the days following the September 11, 1973, coup, a representative of UNHCR, a representative of the World Council of Churches (Charles Harper), and several Chilean ecumenical leaders created CONAR to secure the exit of the large refugee population. Following on the heels of CONAR, COPACHI was created to extend ecumenical support to Chilean nationals. Harper contacted Bishop Helmut Frenz, the head of the Lutheran Church in Chile, who spoke with the auxiliary bishop of Santiago, Fernando Ariztía, who enlisted the support of Raul Cardinal Silva Henriquez, the Archbishop of Santiago. Cardinal Silva Henriquez held a meeting for selected representatives of the Catholic, Orthodox, Lutheran, Methodist, Pentecostal, and Baptist churches and the Jewish community, and COPACHI, officially under the jurisdiction of the archdiocese of Santiago, was set in motion (Lowden 1996, p. 32). The *Comité para la Paz*, then, was not a product of the Catholic Church as institution; rather, it

¹⁸ Estimated numbers of disappeared persons vary widely depending on the source. The Chilean Truth Commission, whose estimates are widely recognized as conservative due to its stringent criteria for accepting cases, reports 1,156 deaths and disappearances from September to December 1973. The CIA reported 11,000 dead between September and November 1973, while the U.S. State Department estimates ranged to 20,000 for the same period (cited in Smith 1982, p. 288). For a comparison and analysis of diverse estimates for Chile, Argentina, Uruguay, and Brazil, see King (1989).

¹⁹ Isolated criticisms of "individual aberrances of power" emerged from 1974 onward; however, the Church did not officially criticize the systematic abuse of power by the regime until mid-1976 (see Smith 1982, chap. 9).

²⁰ The Church's official position of good faith toward the "Christian" military junta should be understood in contrast to its discomfort with the preceding Marxist regime that it perceived as a threat to traditional Catholic values (see Smith 1982, chap. 9).

developed through the personal and professional relations of religious leaders and others who, regardless of denomination, shared certain basic life commitments and values and recognized the threat of the regime to those values.

The preexisting social networks of progressive religious leaders also linked them to leftist politicians, university faculty, social workers, lawyers, and other professionals who participated in the first generation of HROs (Camus 1985; Lowden 1996; Orellana and Hutchison 1991). The Chilean case is unique in the extent to which institutional connections between the Church, Catholic left political parties, labor unions, community organizations, and Catholic universities had facilitated the prior development of personal networks linking individuals from these spheres of society. Personal ties were formed through institutional connections and shared values linking students and faculty at the prestigious Catholic University to progressive leaders within the Church, to political leaders, activists, and party members of the Catholic left. Institutional and personal connections between representatives of the Chilean Catholic Church and progressives in the universities and political parties were facilitated by the Church's official identification with the social reformist principles espoused at Vatican II in 1968 and reaffirmed a few years later at Medillín (Smith 1982).²¹

Many political leaders on the left had studied at Catholic universities, met each other through involvement in student politics, and maintained ties with faculty and priests. These networks, sometimes going back to the 1930s, proved crucial, as the individuals who became involved in human rights activities in the early years of the dictatorship through COPACHI, and later through the Vicaría and FASIC, were for the most part Catholics from outlawed or suspended political parties. According to Smith (1982, p. 334), the "backbone of the original core team" who initiated programs under the auspices of COPACHI "were Catholics formerly associated with leftist parties," such as Movimiento de Acción Popular Unitaria (MAPU) and the Christian Left Party (IC). Some non-Catholics who had been active in the communist or socialist parties also volunteered, along with several priests and nuns, totaling over 300 professional and staff personnel working as lawyers, social workers, physicians, and clerical help (Smith 1982, p. 334). Because of the high risks and the importance of trust and solidarity in the early 'wave' of resistance, participants

²¹ The Church's discomfort with the Allende government was much more a reaction to that government's Marxist ideology than its plans for social reform (see interviews and survey results in Smith [1982]). Mainwaring (1986, p. 170) suggests that "since the late 1950s, the Chilean hierarchy has been one of the most progressive in Latin America."

were recruited exclusively through the personal networks of the core members. According to a psychologist active in FASIC, "The first generation of human rights activists was clearly made up of a network of persons who, if they didn't all know each other personally, had faith in the friends of those they did. Nobody came in off the street ('Nadie llegó desde la calle')" (Lira 1997).

Whether for primarily religious, ethical, or political reasons, some degree of personal commitment and "solidarity" (Fireman and Gamson 1979; Martín-Baró 1983) toward confronting the abuses of the regime characterized the early participants in HROs. Those who began to organize during the first wave, in the midst of unprecedented political repression and state terror, were motivated to do so in part because of moral, personal, and political commitments developed in their life experiences prior to the coup. Preexisting ties and solidarity among certain religious leaders, academics, politicians, and professionals (e.g., lawyers, social workers) facilitated the formation of COPACHL.

As outlined above, recent theories suggest that, to understand why individuals risked their lives by engaging in human rights work in the early to mid-1970s in Chile, it is helpful to consider both participants' self understanding and commitments prior to their involvement and how their "sense of self" is affected by involvement in high-risk collective action (Calhoun 1991, p. 69). Calhoun's (1991, p. 51) suggestion that risks may be borne "not because of the likelihood of success in manifest goals but because participation in a course of action has over time committed one to an identity that would be irretrievably violated by pulling back from risk" is supported by the account of a psychologist working with FASIC who became deeply involved in human rights activism: "In my view, motivations to participate were ethical, political, and very personal. For me, the suffering of the people I was helping was intolerable, the persecution of my students, their disappearance and death still cause me pain today. I believe that one commits oneself to things because of who one is. I believe that I would have lost my own dignity and self-respect if I hadn't done the work I did" (Lira 1997).

Once involved in the daily activity of assisting victims of human rights abuses and their families, saving lives by coordinating asylum for targets of persecution, and helping subjected communities to meet their basic needs, participants became increasingly committed to what Martín-Baró (1983) refers to as "prosocial" collective action and identity, in which commitment to the community or greater good of society outweighs concern for individual needs or satisfactions. As the FASIC psychologist notes: "For many people on the left who were unemployed, it was work. For others, it was a cause, a meaning of life. And if that's how it is, risks lose their importance" (Lira 1997).

The tight personal networks that made up the first generation of HROs were relatively effective and sustainable during the most intense period of repression (1973–76) because they were embedded within extended organizational networks that included the Catholic Church; as a nexus linking immediate social networks of committed individuals to those in need of assistance, the Church as institution played an invaluable role. The “politically neutral” Church organizations provided a degree of symbolic and physical space within which religious leaders, political activists from banned political parties, and concerned professionals could meet and coordinate assistance for victims of persecution or repression.

In addition to organizations explicitly focused on human rights, the Church also facilitated or sponsored a plethora of NGOs, including research institutes such as the Academia de Humanismo Cristiano, which housed professors purged from the universities by the government (Vidal 1986). Along with the Church’s own information gathering and publications (Garretón 1983, p. 179; CEP 1974²²) those of the Academia became an important source of information and “counterinformation” in combating government censorship. Because of the breadth of the Church’s ties to Chilean society, expanded further under the dictatorship as it increasingly served as a “surrogate” for outlawed forms of organization, the Church’s “organizational network offered a unique opportunity to work for human rights” (Smith 1982, p. 334). The particular way that personal networks reflecting the institutional linkages in Chilean society were embedded within the Church partly explains why and how certain individuals managed to organize HROs in the first months after the coup.

The embeddedness of personal networks of committed activists within the domestic and international institutional networks of the Church also accounts in part for Chilean HROs’ ability to inspire and coordinate international resistance to the Pinochet regime. The embedded networks facilitated the development of an elaborate, clandestine web connecting grassroots activists in Chile to high-level international Church officials, to private and government foundations, to Chilean exiles living abroad. Through these channels, many individuals whose lives were threatened under Pinochet were smuggled out of Chile. These networks also helped to sustain political activity of the Chilean exile community trying to influence international opinion of the Pinochet regime. Although human rights activists from Argentina, and to a lesser extent Uruguay, also testified intermittently in international forums, the Chilean activists were exceptional

²² For a fascinating collection exemplifying the type and extent of “counterinformation” being produced within the Church, see CEP (*Centro de Estudios y Publicaciones* 1974).

in their ability to sustain a veritable international political lobby over a considerable period of time.

The Church as institution was also influential in that even before it began to make official its opposition to the military junta, its historical and institutional role in Chilean society was crucial in providing a shield of legitimacy behind which the early HROs were able to coordinate assistance for victims of persecution. The Catholic Church in Chile had been deeply and directly involved in politics since colonial times and in numerous ways remained extremely influential in the moral, social, and political life of Chilean society after the separation of Church and state in 1925 and throughout the decades leading up to the coup in 1973 (Smith 1982).

In its first year in power, the military junta did not wish to risk an open attack on the "humanitarian work" of COPACHI, for fear of risking a confrontation with the Church hierarchy, which might undermine its legitimacy in the eyes of Chileans and in the international community. While not immune to persecution, the work carried out openly under the auspices of COPACHI escaped direct repression, at least for a while. However, toward the end of 1974, COPACHI volunteers and programs began to be targets of government harassment.²³ The implication of a group of priests and nuns linked to COPACHI in procuring asylum for four members of the Movimiento de Izquierda Revolucionaria (MIR) in foreign embassies led to the arrests of several religious members of COPACHI in November of 1975. Pinochet accused COPACHI of providing a front for Marxist-Leninist agitators and, in a letter to the cardinal, he requested that it be disbanded (Smith 1982, p. 318). Although the cardinal acquiesced, formally ending the interdenominational Comité para la Paz, the following month he created a new organization, the Vicariate of Solidarity (Vicaría de la Solidaridad), which essentially took up where COPACHI had left off. According to Smith (1982, p. 318):

While there were some changes in personnel, the new organization continued the same services of COPACHI and was made an integral part of the juridical structures of the Archdiocese of Santiago. While the Church had lost a tactical skirmish with the government, the strategy of the cardinal was shrewd and foresighted. The new Vicariate of Solidarity was more closely tied to the official Church than its predecessor, making it both easier for the bishops to control and harder for the government to smash without directly attacking the core of the Church itself.

The Church as institution provided a "moral shield" for human rights work through its domestic influence as a source of legitimacy and its international symbolic, moral, and political weight. These characteristics of

²³ See Smith's (1982, p. 317) account of "paid spies" attending Church activities.

the Church also made it logistically possible for foreign religious and humanitarian foundations to send money in support of human rights activities without state interference. CONAR and COPACHI benefited from foreign funds from their first moments, as Bishop Frenz arranged for support from the World Council of Churches through his relationship with a member of the council, pastor Charles Harper. Other evangelical and Catholic organizations in Western Europe and the United States, as well as the Inter-American Foundation and the Ford Foundation, beginning in 1974, also made significant contributions (Frühling 1988, p. 149).

The significance of external financing of the Chilean HROs is difficult to overestimate. Smith (1982, p. 325) argues that "none of the new projects begun under the auspices of the Chilean Church since 1973 could have been inaugurated or sustained over time without very considerable outside support." Based on data provided by donating organizations, Smith calculated that between 1974 and 1979 over \$67 million in financial and material assistance was sent to the Chilean Church from Catholic organizations in North America and Western Europe. The Inter-American Foundation contributed approximately \$20 million in grants to Church-sponsored projects. And contributions from North American and Western European Protestant organizations, funneled through the World Council of Churches, totaled approximately \$10 million (1982, p. 326). As Smith points out, these funds dwarfed the Chilean Church's \$4 million in internal resources collected through "tithing campaigns" during this same period. Though the military government attempted on various occasions to obstruct the flow of external funds to the Chilean Catholic Church, it repeatedly backed down in response to international pressure from the Roman Catholic Church, the World Council of Churches, and in the case of Inter-American Development Bank (IDB) funds, members of the U.S. Congress, the U.S. ambassador, and representatives of the IDB (Smith 1982, pp. 327–29). Embedded within influential organizational networks at the international level, the Chilean Catholic Church was able to serve as a funnel for foreign funds for human rights programs, even when the Pinochet regime was willing to risk some domestic confrontation with the Church.

As suggested by resource mobilization approaches in the social movement literature, access to external funding or other resources is frequently crucial for explaining why social movements emerge precisely when they do and are able to sustain themselves over time. It is quite clear to observers and participants in the Chilean human rights movement that foreign funding facilitated both the implementation and the sustainability of human rights programs in Chile and that such funding was both forthcoming and technically possible due largely to Church sponsorship of HROs and their varied activities in Chile (Smith 1982; Lowden 1996).

The "framing" of human rights programs as such also facilitated the moral and technical support of international organizations tied into the international human rights issue network (Sikkink 1993). As indicated earlier, almost immediately after the coup, the UNHCR and the World Council of Churches intervened on behalf of the 10,000–15,000 political refugees residing in Chile. International organizations were thus involved in Chile from the first day after the coup and were not slow to spread the word of massive abuses of human rights. In the first six months after the coup, delegations from Amnesty International and the International Commission of Jurists made visits to Chile to examine the human rights situation. The international publicity and uproar over the abuses committed by the military government immediately following the coup were channeled into an effective political lobby by international HROs and the Chilean exile community, which pressured the U.S. government to pressure the Pinochet regime to halt abuses.²⁴

The combination of foreign funds and international visibility helps to account in part for the sustainability of the first generation of HROs in Chile despite a political opportunity structure that was not conducive to sustained collective action. In the absence of "externally determined" advantageous fluctuations in the structure of political opportunities (Tarrow 1994), resourceful social actors sought out and *created* "opportunity"—through personal networks embedded in historically linked social, political, and institutional networks—where none seemed forthcoming. From there, the emergence of later waves of HROs and the eventual mass mobilization to oppose the continuation of Pinochet in power (1988), and later in the struggle for "truth" and "justice" in the transition to democracy, followed more closely Tarrow's conceptualization of increasing mobilization in response to improvements in the political opportunity structure.

Yet while the political opportunity structure, as captured by the intensity and extent of repression (see figs. A1 and A3 in the appendix) and restrictions on personal liberties and rights,²⁵ did not seem conducive to the emergence of collective action in defense of human rights between 1973 and 1976 in certain respects some Chileans enjoyed a somewhat greater degree of personal freedom in the first years after the coup than in Uruguay or Argentina. Repression peaked in Chile between 1973 and 1976. During and after 1976, arrests and disappearances were more selective than in either Uruguay or Argentina, and there were fewer "errors"

²⁴ A somewhat perverse indicator of the success of the publicity campaign to isolate the Pinochet regime in the international community is that several members of the Argentine armed forces have explained their decision to employ clandestine repressive measures as a means to avoid Pinochet's difficulties with international critics (see Sikkink 1993, p. 423).

²⁵ See n. 10 above.

or "unintended" detentions or murders (in part due to the international publicity mentioned above). For the most part, individuals recognized who the targets of persecution would be and why; this provided some minimal level of day-to-day security, if only for those who fell outside the ascriptive social categories persecuted by the state.

Significantly, in comparative perspective, some social spaces remained open, if constrained, in Chile, that were eliminated either before or upon the military coming to power in Argentina and Uruguay. Though political activity was proscribed, right and center political parties, including the Christian Democrats, were not outlawed until 1977. Some newspapers, including the generally conservative *El Mercurio*, continued to operate without prior censorship, and most social organizations of a nonpolitical nature were allowed to continue to function. Although they were not viable sites for open *political* resistance to the regime, these social spaces allowed people to meet, talk, and develop shared understandings of the situation.

Of course, the relative selectivity of state persecution and the limited availability of certain social spaces would not qualify, by Tarrow's own definition, as aspects of the political opportunity structure that *encouraged* collective action; from Chileans' perspective, these features resulted from sudden changes in the political environment that drastically increased the costs and risks of contentious collective action, which in Tarrow's formulation should *discourage* the emergence of social movements. However, the potential "opportunities" of these features of the political environment may be seen as such when considered in comparative perspective, as will become evident below.

Conclusion

Analyzing the emergence of HROs in Chile within a synthetic theoretical framework suggests that such an approach can improve understanding of the interrelated social processes involved in the emergence of social movements and organizations in high-risk contexts. In lieu of considering either a single "most important variable," or what amounts to a check-off list of hypothetically independent variables operating discretely at distinct levels, this approach sheds light on how several processes discussed in the social movement literature are *interrelated* in particular ways in a historically specific case.

In Chile, certain individuals chose to participate in HROs despite the risks because of moral, personal, religious, social, or political commitments developed prior to the coup and because of preexisting personal connections linking them to others with similar commitments. Once involved, their personal and collective sense of commitment was increased through

the experience of participation. These face-to-face networks, which reflected the institutional networks between the Catholic Church, Catholic University, Catholic left political parties, labor unions, and professional organizations, were embedded within the broader institutional networks of the Church and transnational NGO and religious networks. These networks permitted the rapid emergence and sustainability of HROs during the height of repression by facilitating almost immediate access to international funding and publicity and by providing the symbolic and moral legitimacy and the physical and sociopolitical space from which to coordinate programs in defense of human rights from the first moments after the coup.

URUGUAY

The Absence of HROs during the Height of Repression

Uruguay presents a striking contrast to Chile. Although the *autogolpe* in Uruguay occurred the same year as the Chilean coup,²⁶ the first and only HRO to operate as such under the dictatorship did not emerge until 1981 (Frühling, Alberti, and Portales 1989, p. 262; Weschler 1990, p. 154).²⁷ What requires explanation in the Uruguayan case, then, is the *absence* of HROs in the first seven years under the military regime.²⁸

The Chilean case suggests the importance of access to resources such as funding, information, and sociopolitical or symbolic space, to account for the *timing* of emergence and sustainability of HROs in high-risk contexts. Additionally, it reveals that participation in collective action in high-risk contexts depends on particular types of face-to-face networks, the way in which immediate social networks are embedded within broader institutional networks, and the forging of increasingly committed self and collective identities in the process of participation. In the Chilean case, a particular configuration of these factors facilitated a rapid and sustained response by particular actors whose activities created additional space for subsequent waves of activists, whose mobilization corresponded to relative improvements in the political opportunity structure. The Uruguayan case provides the opportunity to assess how well this same combination

²⁶ The Uruguayan armed forces denied that a coup had taken place in Uruguay in 1973, when the National Congress was dissolved, because the elected president, Bordaberry, retained the office of chief executive. They place the date of the coup three years later, on June 12, 1976, when Bordaberry was forced to resign.

²⁷ See n. 16 above.

²⁸ To my knowledge, no one has explored why HROs formed so much later in Uruguay than in Chile during this period. Of course, the absence of collective action is the expected outcome for those who follow Mancur Olson (1965).

of factors accounts for the *absence* of the outcome they explain in the Chilean case.

Uruguayan and Chilean societies were similar in several important respects prior to the early 1970s. Each enjoyed a highly developed political party system, national labor unions, a large urban population, a prestigious education system, and a long electoral tradition with only a few brief interruptions.²⁹ In one important way, however, Chile and Uruguay could scarcely have been more different. The secular nature of Uruguayan society contrasted sharply with the permeation of religion and the Church in most all spheres of Chilean life, from education to politics. Uruguay is widely recognized as an extremely secular society, in which the Church has little influence outside a clearly demarcated and limited "religious" sphere. This contrasts with the Chilean Catholic Church's long history of strong political and social influence, which in recent decades had been exerted increasingly toward social reform. The Church in Uruguay has been "historically weak" and "has never assumed the defense of the persecuted or oppressed" (Gauding 1991, p. 86). The Church in Uruguay also lacked strong historical connections to the major political parties or labor organizations. The largest party in Uruguay, the Colorados, was ardently secular, even anticlerical, in contrast to Chile's most important political party, the Christian Democrats. When compared with the Chilean case, a theoretically informed consideration of this feature of Uruguayan society provides several insights that help to account for the absence of HROs in Uruguay until the political climate shifted following the plebiscite in 1980.

First, and perhaps reflecting the isolated and limited role of the Church in national political life, neither the Catholic Church nor other religious institutions were able to provide the kind of protective moral "shield" from direct persecution that COPACHI depended upon during the first months of the Pinochet regime. Whereas COPACHI's affiliation with the Church warded off (at least for a while) direct persecution by the military regime,³⁰ the Uruguayan armed forces openly confronted the Church *as institution* in their first years in power. In the years of armed struggle between the armed forces and *Tupamaro* guerrillas, which lead up to the *autogolpe*, several prominent members of the Uruguayan Bishop's Conference, including the archbishop of Montevideo, espoused increasingly progressive

²⁹ Due to its democratic legacy, large middle class, and commitment to a welfare state, Uruguay had earned the international reputation of the "Switzerland of Latin America." In specialists' rankings of "political development" in Latin America prior to the seventies, Uruguay is generally ranked first among Latin American countries, with Chile following close behind. See, e.g., Fitzgibbon (1954)

³⁰ Smith (1982, p. 312) suggests that regime's initial tolerance of COPACHI was a trade off for the acquiescence of the Church hierarchy during the first year after the coup.

views reflecting support of the principles of Vatican II and Medellín.³¹ In 1969, and again in 1972, the Bishop's Conference issued letters that, despite their overall neutral tone and calls for "reconciliation," offered subtle critiques of the increasing use of torture and unjust imprisonment (Kaufman 1979, p. 45; Inter-Church Committee on Human Rights in Latin America 1978, p. 15). President Bordaberry's response to the 1972 letter essentially dismissed the bishops' concerns. In part, he replied: "In this struggle, conventional standards are not applicable. . . . Information is a decisive factor, it is the basis of success. . . . Information is obtained in some instances spontaneously . . . and in others after rigorous interrogations. I defend the rigor and the severity of interrogations, which avoid bloodshed and deaths in this war, and which make possible bloodless victories."³²

Following the *autogolpe* in 1973, which sparked protests by unions and student organizations and resulted in massive arrests, the Church maintained a position of official silence (Rama 1987, p. 175). But as Uruguayans were soon to learn, neutrality was not an option in the generals' Uruguay. The Church's official silence was interpreted as a failure to cooperate with the military government, and the Church was considered infiltrated by communism "whose ruinous villainous and treasonous actions must be once and forever expurgated" (General Forteza in *La Opinión*, September 29, 1973; quoted in Kaufman 1976, p. 46). According to General Forteza, international communism "has reached the Church itself, violating in this institution the rights and obligations that the State has granted to the different religions" (Forteza in *La Opinión*, September 29, 1973; quoted in Kaufman 1976, p. 54).

In contrast to the church-state relationship under military rule in Chile, in Uruguay, not only were the Church and other religious institutions perceived as subordinate to the state, the Church was accused of violating the conditions of its existence as determined by the state. This is clear in the open verbal attacks on the Church as institution,³³ as well as in the

³¹ Several prominent members of the Church hierarchy also defended the traditional, conservative role of the Church; overall, the Uruguayan Church was (by default) more progressive than in Argentina but much less so than in Chile. Mainwaring (1986, p. 115) suggests that Argentina and Uruguay are exceptional among Latin American Roman Catholic Churches in that they failed to become more progressive in the repressive situations of the 1970s.

³² *Ahora*, June 16, 1972 (cited in Inter-Church Committee on Human Rights in Latin America 1978, p. 15).

³³ While the Pinochet government criticized particular members of the Church on numerous occasions, it avoided verbal attacks on the legitimacy of the Church as institution.

persecution and arrests of numerous religious functionaries. In addition to the arrest, torture, and, in some cases, death of individual Methodists and members of Catholic orders (such as Jesuits and Dominicans), the regime also created an official commission to investigate Catholic Church activities (Inter-Church Committee on Human Rights in Latin America 1978, p. 51). According to the former auxiliary bishop of Montevideo, Andrés Rubio: "The Uruguayan police tentatively watches the Catholic Church, controls and watches the content of sermons in the churches and investigates the text of the material circulated; several parishes and houses of clergymen have been subjected to searches and some priests have been arrested" (in the Mexican newspaper *Excelsior*, June 21, 1975; quoted in Kaufman 1979, p. 81).

Largely due to its relative isolation from the political process and historically marginal position in other spheres of Uruguayan society, the generals did not feel obligated even to feign respect for the Church. With members and leaders subject to arrest and imprisonment, the Church as institution could not provide "protected spaces" to the extent that it could in Chile, nor could it effectively perform the role played by the Chilean Church of collecting and disseminating "counterinformation." In Chile, the historical, political, and symbolic importance of the Church as institution forced the military junta to adopt a more cautious stance toward the Church, allowing it to become the single most important locus of resistance and moral opposition to the military regime. In Uruguay, the Church was much more easily controlled and repressed by the generals; it may well have been the "weakest of all political pressure groups in the country" (Kaufman 1979, p. 81).

While numerous sectors of "civil society" in precoup Uruguay were highly organized, the *particular type* of cross-sectoral personal networks that made such a difference in Chile were lacking. The contrast between the relationship of the church and the political left in Chile and Uruguay is particularly striking. Until quite recently, the left was historically insignificant as a political force in Uruguay, and the Christian left even more so. The two major political parties in Uruguay were both secular; the traditional Colorado and Blanco parties had monopolized political control through an elaborate electoral system and power-sharing arrangement that guaranteed the "loser" (for most of this century, the Blancos) one-third of the seats in legislative bodies. In the 1966 elections, the Christian Democratic Party won a mere 3% of the votes, and all leftist parties together accounted for only about 10% of the votes (Franco 1984, p. 95). With the formation of the Frente Amplio, a leftist coalition including the Socialist Party, the Frente Izquierda, the Christian Democratic Party, and the Union Popular, among other (even) smaller parties, the combined

votes for the left in the 1971 elections rose to 18.28%, but the Christian Democratic Party still claimed only 3.86% of the electorate (Franco 1984, p. 95).

Whereas in Uruguay no significant Christian Democratic Party had developed prior to the *autogolpe*, in Chile before the coup, the Christian Democrats had governed from 1964 to 1970 and remained the country's single most powerful electoral force; the Christian left was an important if small part of the Allende coalition. In Chilean municipal and parliamentary elections since 1963, the Christian Democrats had received more votes than any other single party, and combined with the socialist and communist parties, claimed over 50% of votes in all elections through 1973 (Bravo Lira 1978, p. 203). In contrast to the 3% of votes won by the Christian Democrats in Uruguay in 1971, the Chilean Christian Democrats earned 25.72% of total votes in municipal elections the same year and 29.12% in the parliamentary elections two years later (Bravo Lira 1978).

The comparatively low support for the Christian Democratic Party in Uruguay, yet another indicator of the secular nature of Uruguayan society, reflects the historically low level of interconnectedness and ideological affinity between the Church, political parties, and labor organizations.³⁴ Institutional linkages between the Church and the most prestigious universities, where the political and professional elites were predominantly educated, were also lacking. In contrast to Chile, where Catholic and secular higher education are both esteemed, in Uruguay the production of future political leaders, lawyers, academics, psychologists, doctors, and social workers occurred almost exclusively in public, secular universities. Political leaders were not influenced by Catholic education, nor had they participated together along with other would-be professionals in Catholic youth groups, as was often the case in Chile since the 1920s. Hence, another crucial link in the type of face-to-face networks that were so consequential in Chile was missing in Uruguay. The lack of institutional linkages between Church, universities, and political parties in Uruguay suggests that even if the Church had been able and willing to provide a relatively secure space for coordinating efforts in defense of human rights following the *autogolpe*, it would have had few "secure" channels through

³⁴ Writing in 1964, Goran G. Lindahl commented: "The Catholic party, the Unión Cívica, has never gained much support, not only because Uruguay is probably the most atheistic country in Latin America, but possible also because Batlle [the country's most important political leader in the early 20th century] was strongly anti-Catholic: he said that the Catholic religion, like all other religions, was filthy. Today Batlle's son, who runs his old newspaper, *El Día*, goes on writing 'god' without a capital letter. The Unión Cívica has not even been able to gather many of the Catholic voters, most of whom seem to vote for the Nationalists" (Lindahl 1964, p. 450).

which to recruit participants from other sectors of society without the risk of infiltration from security forces.³⁵

The Church's weak position in Uruguayan society helps to explain why the types of individuals who became involved through personal networks in the first generation of HROs in Chile—religious leaders and functionaries, party members from the Christian left, academics, lawyers, and social workers—did not create HROs in Uruguay in the face of military dictatorship. However, it only partly accounts for why no organizations emerged at all: Why did no other social actors, perhaps embedded in distinct types of social networks from those that linked the Chilean activists, act collectively to form HROs prior to the 1980s in Uruguay?

Though a reasonable question, the answer is rather obvious. In the Chilean case, *particular types* of social networks, embedded in the multidimensional organizational networks of the Church, were *key* for launching a quick and sustained collective effort in defense of human rights because other potential spaces, particularly political parties, the labor movement, the universities, and professional associations, were restricted following the coup in 1973. This was true to an even greater extent in Uruguay, where traditional spaces for both political and nonpolitical organization were rapidly and systematically eliminated or brought under military control. Uruguayans experienced "the systematic destruction of all the spaces (*ambitos*) that surrounded the State, or that developed autonomously and that could effectively, or eventually, contend for power, information, or cultural production, as well as [military] intervention or mediation in all intermediate forms of social organization that could possibly, regardless of intentions or objectives, become spaces of refuge for the persecuted or those excluded from power" (Rama 1987, pp. 169–70).

Consideration of the nature and extent of repression in Uruguay can largely account for why other social actors, embedded in diverse secular social networks, failed to coordinate and sustain HROs prior to 1981. To an even greater extent than the other countries under consideration, Uruguayan society was thoroughly and deeply penetrated by the monitoring and repressive apparatus of the military state.

Both geography and demography (over 50% of the population lives in the capital city of Montevideo) contributed to the armed forces' social control capabilities. According to the Lawyers Committee for International Human Rights, "Uruguay was the closest approximation in South America of the Orwellian totalitarian state. A small and demographically

³⁵ This was a danger even for the Chileans; however, in an interview with the author, a psychologist with FASIC commented that despite three separate attempts by security forces to infiltrate the Vicaría, they never succeeded because (in a situation where "nadie llegó desde la calle") they were always identified as outsiders.

homogeneous country, without internal geographical barriers, Uruguay became a laboratory for the national security state" (Lawyers Committee for International Human Rights 1985, p. 51).⁴⁶ An important difference from Chile is that the foundations and infrastructural skeleton of this "laboratory" were already in place prior to the disbanding of the Uruguayan congress in 1973. In Uruguay, as in Argentina, the armed forces had confronted real and violent "internal enemies" (the Tupamaros and Montoneros, respectively) before coming to power. As a consequence, the security forces were already organized and mobilized for waging a broadened "war against subversion,"⁴⁷ and important "wartime" legal restrictions on the population were already in effect when the military took control of the state.⁴⁸ In contrast to Argentina, the armed forces in Uruguay eschewed a clandestine "dirty war" strategy of massive "disappearance" and physical elimination of "subversives" (see table A1 in the appendix); instead, they enforced severe legal restrictions of individual freedoms and "legally" subjected thousands of Uruguayans to a prison system designed to systematically destroy individual personality through physical and psychological torture (Weschler 1990, pp. 131–47; Amnesty International 1983).⁴⁹

In contrast to Chile, in Uruguay most potential spaces for organized resistance had already been restricted or eliminated before 1973. Although the repression in the first moments after the coup in Chile was severe, in comparative perspective, the Chilean security forces were not as prepared "logistically" (e.g., clandestine torture centers were not already established and fully functioning, the DINA [intelligence agency] had only just been created) to systematically persecute and paralyze all opposition from every sector of civil society as were their counterparts in Uruguay. When compared to Uruguay, it seems plausible that the confusion that followed the coup in Chile—both within civil society as well as within the armed

⁴⁶ The extent of social control is illustrated by the creation of a computerized classification system designating all public employees as "A" for politically "clean," "B" if involved in "some dissident political activity but capable of rehabilitation," or "C"—"banned." This system resulted in massive firings and facilitated identification of "inactive subversives."

⁴⁷ Evidence that security forces were organized and in operation prior to 1973 is provided by a report issued in 1969 by a senate named Special Investigative Commission of the Violation of Human Rights, Torture of Prisoners and Conditions of Detention Afflicting Human Dignity. The report includes testimonies of torture victims, coroners, public defenders, and lawyers. See *Acts of the Uruguayan Senate-II-7, Information documentaire d'Amérique Latine* (INDAL), Belgium: 247–82 (cited in Inter-Church Committee on Human Rights in Latin America 1978, p. 4).

⁴⁸ See n. 10 above.

⁴⁹ The main prison under the military was called, ironically, *Libertad*. Prisons were methodically designed, with the assistance of psychologists, to "demolish the mental, emotional and moral integrity of their inmate populations" (Weschler 1990, p. 131).

forces—regarding what had happened and what would happen in the months ahead left some spaces open and conducive to the coordination of collective resistance immediately following the coup. In short, the political opportunity structure in Uruguay was even less favorable for the creation of HROs than in Chile in 1973.

Observers of the repression in Uruguay emphasize that the breadth and depth of military involvement in all spheres of life, and the *extent* of state repression, was unmatched by other Southern Cone regimes. Though the statistical comparisons of deaths and disappearance under the Southern Cone dictatorships make the Uruguayan dictatorship seem relatively benign (see appendix table A1), these numbers are misleading as indicators of the extent of repression because the generals in Uruguay opted for a different repressive strategy: massive arrests, torture, prolonged imprisonment, and intervention in all spheres of life, public and private (*Servicio Paz y Justicia* 1989, chap. 3).⁴⁰ Under the military government, Uruguay had the highest concentration of political prisoners in the world. In 1976, Amnesty International estimated that “one in every 500 inhabitants of Uruguay was in prison for political reasons and that one in every fifty citizens had been through a period of imprisonment, which for many included interrogation and torture” (Amnesty International 1983, p. 1). The Lawyers Committee for International Human Rights (1985, p. 52) estimates that one in every 47 Uruguayans was “subjected to some form of repression during the dictatorship, whether in the form of house arrest, torture, beatings or a house raid.” These numbers earned Uruguay the reputation of the “great lockup.”⁴¹ And though arrests generally targeted certain sectors of society (such as members of political parties, unions, and student organizations), due to the broad definition and interpretation of what constituted grounds for arrest in the interests of “national security,” no one was immune.⁴² The creation and persecution of “thought crimes” through laws that prescribed prison sentences for “the *intention* to commit a crime” or “to damage the honor of the armed forces” or otherwise threaten the nation, meant that anyone could be arrested without any

⁴⁰ The military intervened directly in all public administrations, universities, unions, and professional associations.

⁴¹ The extraordinarily high numbers of exiles during this period is another indicator of the extent of repression. An estimated 300,000 people fled Uruguay in this period, including an estimated one-third of the population between the ages of 20 and 35 (Lawyers Committee for International Human Rights 1985, p. 4). For comparative estimates see Inter-Church Committee on Human Rights in Latin America 1978, p. 9; King 1989.

⁴² Victims included priests, nuns, high school and university students, teachers, professors, journalists, lawyers, political party members, union leaders and members, and even medical workers (*Servicio Paz y Justicia* 1989, chap. 5).

legal recourse. Lawyers who defended those accused of such crimes were considered to be ideologically sympathetic to their clients and therefore also subject to arrest.⁴³ Hence, the Uruguayan generals' strategy effectively instilled a paralyzing culture of fear (Corradi et al. 1992) throughout society, and deep intervention by the military in all spheres of life closed off potential spaces for the emergence of HROs.

When the defeat of the generals' constitutional referendum in 1980 created temporary openings in the political opportunity structure, one HRO did emerge. Servicio Paz y Justicia (SERPAJ)-Uruguay, created in 1981, "was the first organization devoted to work on behalf of the victims of repression and poverty to be established in Uruguay since the advent of the dictatorship" (Frühling et al. 1989, p. 262). In the same spirit as its sister organizations operating in Argentina since 1974 and Chile since 1977, SERPAJ-Uruguay sought to raise consciousness about human rights abuses (grassroots education programs), document human rights abuses committed by the military government, provide economic assistance for medical treatment for victims, and assist returned exiles. It also coordinated groups of family members of disappeared and political prisoners. However, the Christian-humanist organization, led by Perez Aguirre (who was arrested several times and released largely in response to international pressure; see Weschler 1990, p. 155) was severely limited in its efforts by a lack of external funding.

Here, another crucial role played by the Church in Chile again provides an important contrast; the Chilean Catholic Church made it logistically possible for international foundations to send financial support for humanitarian projects. In Uruguay, there was no institutional "funnel" through which funds could be "anonymously" received and diffused. According to a representative of a Swedish ecumenical NGO that funded several human rights programs in Latin America: "SERPAJ received a series of offers of economical support from foreign sources. However, it had to refuse in almost all cases because there was no way to introduce the resources into the country. Several people returning to Uruguay offered to take part and deliver money personally, but the founders [of SERPAJ] did not want to risk their safety" (Gauding 1991, p. 87). A member of SERPAJ-Uruguay commented, "I remember that when I was outside Uruguay people used to ask me, 'how can we send money to Uruguay?' There wasn't a single person who dared to receive a check, to give to a

⁴³ The fact that lawyers in Uruguay who defended those accused by the regime were themselves criminalized (which occurred to some extent in Argentina as well) undermined any possibility that the legal system might provide some recourse for the persecuted. On lawyers under the military in Uruguay, see Kaufman (1979, p. 77).

family member of a political prisoner, for fear of being sent to La Libertad prison" (as quoted in Gauding 1991, p. 87).

Conclusion

Unlike the other cases examined in this article, the Uruguayan case conforms quite neatly with sociological and commonsense expectations that severe state repression corresponds with generalized demobilization (Tarrow 1994; Tilly 1978). Comparative analysis suggests that this is due to the inability of the Church in Uruguay to play a role parallel to that played by the Church in Chile—as "moral umbrella" and as a funnel to capture a large and constant flow of resources—*combined with the regime's repressive strategy*, which effectively obstructed alternative sites of opposition. The combination of theoretical approaches applied to the Chilean case to explain the rapid emergence and sustainability of HROs under Pinochet can also account for the delayed emergence and nonsustainability of HROs in Uruguay. The Argentine case provides the opportunity to assess the utility of this approach to explain a more "intermediate" outcome.

ARGENTINA

An Intermediate Case

Attesting to the value of incorporating multiple cases in comparative research, the Argentine case suggests a challenge to the conclusions drawn from comparison of the Chilean and Uruguayan cases: without the support of the Catholic Church as institution, HROs still emerged under the harshest years of military rule in Argentina. In fact, not only did the Church—a powerful influence in Argentine politics and society—fail to support programs to defend human rights and assist victims of persecution, on numerous occasions it publicly voiced its support for the dictatorship.⁴⁴ Hence, while the Church in Uruguay seemed (at least somewhat)

⁴⁴ As previously mentioned, the reasons for the Argentine Catholic Church's complicity with the dictatorship are beyond the scope of this article. For a concise historical-institutional explanation, see Mignone (1986, pp. 72–92). Illustrative of the Church's complicity, Cardinal Aramburu publicly denied that the disappeared existed and maintained that position despite increasing evidence presented to the bishops by victims' families. According to Lowden (1996, p. 18): "The Episcopal Conference was even silent in the face of the persecution of clergy involved in human rights work, itself of unprecedented proportions: sixteen priests were murdered or 'disappeared' and two bishops died under highly suspicious circumstances" (see also Mignone 1988, chap. 2, chap. 8).

willing, but unable, to provide space and support for HROs and the Church in Chile proved both willing and able, the Church in Argentina was able but far from willing. This meant that neither the semiprotective "shield" of symbolic and moral legitimacy nor the multilevel organizational networks of the Catholic Church were available to human rights activists in Argentina. Despite the absence of conditions that were crucial for the emergence of Chilean HROs immediately following the coup, and that account for the absence of HROs in Uruguay until 1981, HROs arose in Argentina during the worst years of repression (1976–79). In the midst of massive and horrific repression in the mid- to late 1970s (see fig. A2 and table A1), and without the protective cover of the Church, why were certain individuals able to organize and act collectively in defiance of the military junta in Argentina? Did these organizations differ in significant ways from those that first emerged in Chile with Church support?

Much like in Chile, the Argentine Catholic Church is a powerful influence in Argentine society and is traditionally a major source of legitimacy for some political leaders.⁴⁵ In contrast, however, the two major political parties in Argentina, the Peronistas and the Radicals, were aggressively secular, if not anticlerical. As in Uruguay, there was no significant Christian Democratic Party and no generational or institutional connections of the Church to major labor, student, and professional associations. Thus, the embedded networks that connected activists to the Chilean Church were largely lacking in Argentina. The complicity of the church with the dictatorship in Argentina closed off what could have been a decisive source of opposition to the abuses of the military junta (Mignone [1986] 1988 p. 21). In stark contrast to the Chilean Church, the Argentine Church aligned itself with the military junta, granting its support and legitimacy to the imposed regime. As in Chile and Uruguay, divergent opinions and attitudes coexisted within the Church hierarchy; but in Argentina, radical and progressive members of the hierarchy were a marginalized minority.⁴⁶

⁴⁵ In fact, in the 1970s, there was still no formal separation of Church and state in Argentina, and the constitution stipulated that the president must be Catholic.

⁴⁶ An organization of radical priests critical of the hierarchy emerged in 1968—the Movement of Priests for the Third World—but it was marginal and, as it earned a reputation of being a socialist organization, fell prey to repression under the 1976 junta (Mainwaring 1986, p. 168). Indicative of the minimal institutional penetration of progressive ideas in the Argentine Catholic Church, Mainwaring suggests that moderates and conservatives within the Argentine hierarchy were unwilling to defend progressive priests or bishops targeted by repression. This is supported in documents presented by Mignone (1988 chap. 8) revealing the collusion of the junta and the conservative hierarchy in the persecution of progressive sectors of the Church (the exact opposite occurred in Chile, where attacks on progressive religious functionaries caused the Church to finally take an official stand in opposition to the regime) (Smith 1982, chap. 9).

The historically conservative hierarchy—which had largely ignored, if not opposed, the social reformist tendencies stemming from Vatican II and Medillín—refused to denounce mounting evidence of human rights abuses, and the majority of churches closed their doors on family members of victims who came to plea for help (Mignone 1988, chap. 2). According to Mignone (1988, p. 19): “The Argentine episcopacy is made up of more than eighty active prelates, including heads of dioceses, auxiliary bishops, and military bishops. Only four of them took a stand of open denunciation of the human rights violations committed by the terrorist regime.” The types of spaces, opportunities, and resources provided by the Church to activists in Chile were not forthcoming to their counterparts from the Church in Argentina.

With the Church doors closed and traditional spaces for organizing and coordinating collective action restricted or eliminated through legal and extralegal measures, HROs in Argentina emerged primarily through non-institutional channels. In addition to the *Liga Argentina por los Derechos Humanos*, a civil libertarian group dating to 1937 with unofficial links to the Communist Party, three other organizations came onto the scene *prior* to the 1976 coup that established the military dictatorship.⁴⁷ While these HROs emerged in response to repression, they did not initially confront the extent and intensity of repression imposed after 1976 by the military junta. Ecumenical leaders were active in forming all three of them, and two of them operated through personal relations of religious leaders to their parishes. This suggests an important contrast to Uruguay, where (for reasons discussed above) neither the Church—nor much less any other religious organizations—were fertile grounds for the emergence of HROs. Rising numbers of human rights abuses preceded the coup in Argentina because, as in Uruguay, sustained armed conflict between urban guerrillas and the armed forces had led to the erosion of legal protections of individual liberties and rights and to the escalation and institutionalization of extralegal forms of repression.⁴⁸ In response, in 1974, an Argentine section

⁴⁷ See n. 16 above

⁴⁸ Like the Tupamaros in Uruguay, the Montoneros in Argentina provided justification for increasing military control over Argentine society. While neither the Tupamaros nor the Montoneros posed anywhere near the military threat claimed by the Uruguayan and Argentine armed forces to justify the tactics of their respective “wars against subversion,” they were not “imaginary” enemies either. It is easy, and tempting, to overlook the provocative actions of these urban guerrilla movements in the face of evidence of the horrific abuses committed by the military regimes. But if their repressive actions remain unjustifiable, when the Montoneros were so bold as to wage armed attacks on military barracks, it was not totally unreasonable for the armed forces to speak of an “internal war.” In contrast to Uruguay and Argentina, the Chilean armed forces did not face a significant armed threat before taking power, though smaller armed guerrilla movements later arose in response to the military regime.

of SERPAJ began providing assistance to grassroots sectors suffering effects of the heightened repression. The work of SERPAJ focused primarily on popular education and general assistance for marginalized communities, along with support for the creation of other HROs.

Some ecumenical leaders in Argentina who participated in SERPAJ grew frustrated by its "passive" stance in the face of the suffering they observed among their parishes and broke off to form the *Movimiento Ecu-ménico por los Derechos Humanos* (MEDH) in February 1976, on the eve of the coup. MEDH, which included dissident Catholic clergy and Protestant leaders among its founders, provided direct assistance to victims of human rights abuses and their families. As was the case for CO-PACHI in Chile, a combination of personal networks and external resources was crucial for the formation and activity of MEDH. In an interview, a member of MEDH recalled how "together with some friends, we considered forming a net of solidarity in different parishes to resolve concrete cases" (quoted in Gauding 1991, p. 103). Through mutual friends, members of MEDH were introduced to a representative of Diakonia, a Swedish NGO that offered financial assistance. Face-to-face relations were essential, as suggested by another member of MEDH: "In that period, it was impossible to work within any institutional structures. The Swedes ran the same risks as we did. We had to trust each other [*Teníamos que confiar el uno en el otro*]" (quoted in Gauding 1991, p. 103). Along with these two religious-based organizations, the *Asamblea Permanente por los Derechos Humanos* was formed by individual religious leaders, lawyers, academics, politicians, and other professionals as an alternative to the Liga with its reputed communist sympathies.⁴⁹ The primary activity of the APDH was the collection and systematic documentation of human rights abuses and disappearances. As in Chile, the prior religious, moral, and political commitments of certain individuals embedded in particular face-to-face networks facilitated the formation of HROs, despite severe repression in the mid-1970s in Argentina.

However, there are important differences in the early HROs that emerged in Chile and Argentina. If scholarly discussions of the human rights movement in Argentina rarely focus on the work of these organizations in the first years after the coup, it is only partly because the spotlight was diverted by the emergence of the *Madres de Plaza de Mayo* in 1977, followed by the *Abuelas de Plaza de Mayo* shortly thereafter. It is impressive that without the support of the Church—and at times with its explicit condemnation—HROs formed prior to the coup (before repression dra-

⁴⁹ Like MEDH, the APDH was created through personal networks, since, as Mignone explains (1991, p. 101), individuals formed APDH "without representing—and often against the wishes—of the collectivities to which they belonged."

matically increased) and managed to provide invaluable assistance to victims of human rights abuses and their family members. However, in comparative perspective, their development and activities "were severely limited due to the lack of support from the Catholic Church" (Frühling 1988, p. 161).

The absence of Church support affected these organizations in a number of overlapping ways. For example, in contrast to COPACHI, in their first moments, the activities of both MEDH and APDH were limited by lack of resources (Gauding 1991; Mignone 1991, p. 102). Eventually, both organizations benefited from connections to the World Council of Churches (Brysk 1994, p. 51; Mignone 1991, p. 102), which provided a degree of financial support and, particularly for MEDH, served as an alternative source of moral legitimacy. However, in predominantly (and officially) Catholic Argentina, the moral authority of the World Council of Churches could not offer the degree of protection from direct persecution that programs under the auspices of the Church in Chile enjoyed in the months after the 1973 coup.

During the mid- to late 1970s, the religious leaders involved in SERPAJ and MEDH and the heterogeneous members and leadership involved in APDH provided assistance to victims of repression at great risk to themselves. Brysk's (1994, p. 56) recounting of several cases of persecution of human rights activists is illustrative:

Repression of the human rights movement touched every organization, affecting both the leadership and grass-roots membership. Many members of the original leadership of *Las Madres* "disappeared," while the *Movimiento Ecu  nico* lost two nuns, several priests, and a Protestant minister. The co-founder of the *Asamblea* . . . was kidnapped, tortured, and imprisoned for several years. . . . Several *Liga* lawyers disappeared,⁸⁰ and a secretary of the *Familiares* was kidnapped, tortured, and forced to give false statements to the press denying her disappearance and alleging connections to guerrilla forces. Rank-and-file members of *Las Madres* were arrested repeatedly following demonstrations. . . . The offices of *Asamblea*, *CELS*, *La Liga*, and *Movimiento Ecu  nico* were raided.

Mignone (1991, p. 104) recounts how the meeting places of APDH were bombed on numerous occasions. This repression prevented these organizations from providing nearly the extent of support or protection from state violence that programs under the auspices of the Church in Chile were able to provide in the first months under Pinochet. The absence of the Church umbrella also deprived Argentine HROs of the social and

⁸⁰ Again, this illustrates the less favorable political opportunity structure in Argentina than in Chile, where lawyers were exiled and jailed but none were killed or disappeared.

organizational networks that fostered the emergence, growth, and sustainability of the HROs in Chile. Overall, according to Fröhling, in the first few years under the military junta "none of the Argentine human rights organizations reached the level of development or extent of human rights programs comparable to what was occurring in Chile at that time" (Fröhling 1988, p. 162). Theoretically informed comparison of Chile, Uruguay, and Argentina, synthetically drawing on various strands of social movement theory, thus helps to explain both why these HROs emerged at all in Argentina and why they were relatively less effective than their counterparts in Chile in the early years after the coup.

Following the coup in March 1976, the Argentine military's chosen strategy in its "war against subversion" created a new category of social actor, as had occurred in Chile after 1973: relatives of the disappeared. This "ascribed identity" became the basis for the formation of three new HROs between 1976 and 1977, the *Madres de Plaza de Mayo*, the *Abuelas de Plaza de Mayo*, and the *Familiares de Desaparecidos y Detenidos por Razones Políticas*. Accounts of the emergence of these organizations invariably emphasize the extraordinary extent and cruel nature of repression under the military junta. Attempting to avoid the international criticism that plagued Pinochet, the armed forces in Argentina opted for a policy of clandestine state terror combined with official denial. Brysk (1994, pp. 36–37) describes the characteristic technique of disappearance: "Disappearance' involved kidnapping of unarmed citizens (usually in the middle of the night, from their family homes) by a gang of armed men, followed by forced removal of the victims to clandestine detention centers, extensive torture, and mistreatment, and (almost always) murder. . . . Although the kidnappers usually sought a specific person, other family members or visitors often "disappeared" in lieu of or in addition to the intended victim."⁵¹

This strategy effectively instilled a culture of fear in Argentine society with all its paralyzing effects (Corradi et al. 1992); the majority of the population did their best to live day to day by trying to ignore or deny what was going on around them. The Argentine armed forces were determined to exterminate "subversive cancers" from the body of the nation, and to do so, they targeted vast numbers of Argentine citizens. In the words of Brigadier General Ibérico Manuel Saint-Jean, former governor of Buenos Aires: "First, we will kill all the subversives; then we will kill their collaborators; then their sympathizers; then the indifferent and, finally, the timid" (Comisión Argentina por los Derechos Humanos 1997,

⁵¹ Eduardo Duhalde (1983, p. 146) discusses an internal military document from 1978 that places the margin of error of the disappearance campaign at "no more than 25 percent" (approximately 2,500 people).

p. 13).⁵² The war against subversion in Argentina was not only a physical battle, it was also moral crusade. In response to a Mexican journalist seeking information on the fate of a woman confined to a wheelchair who had been detained, General Videla explained: "A terrorist is not only someone who kills with a gun or plants bombs, but anyone who encourages their use by others through ideas contrary to our Western, Christian civilization."⁵³ For this reason, the military government enacted sweeping censorship laws and even took the trouble to individually ban thousands of books, songs, and films, among them *"El Principito"* (The little prince), in its ninety-fifth edition in Argentina (Garcia 1995).

In these extremely repressive conditions, the founders of Madres de Plaza de Mayo met while searching for information about their disappeared children in government offices. Beginning in April 1977 a small group of these women decided to engage in symbolic protest against the regime by marching in the central public space of the nation, the Plaza de Mayo. Despite government attempts to crush the organization (nine of the original founders were "disappeared" after their meetings were infiltrated by an undercover military officer) and continuous persecution (members were frequently arrested following demonstrations), the Madres and Abuelas continued their weekly marches in the plaza, attracting international press coverage for their vigils. Under such an unfavorable political opportunity structure, why did these women risk their own security to protest the disappearance of their children and grandchildren?

Precisely because the decision by a small group of politically inexperienced women to stand up to the military junta seems so extraordinary and in many ways incomprehensible,⁵⁴ numerous authors have grappled with this question. Some accounts suggest that these women were able to confront the regime precisely because their claims—and the "identity" upon which they were based (motherhood)—were apolitical and were voiced in terms that challenged the regime's own discourse on "defense of the family." However, the disappearance of the founders seems to challenge the thesis that motherhood provided a protected space from which to launch symbolic protest. Other authors suggest that Las Madres protested the disappearance of their children because their sense of self, as

⁵² Quoted in Comisión Argentina por los Derechos Humanos (1977, p. 13; Rock 1985, p. 444, n. 33; Camps 1983, p. 63) David Rock reports that Saint-Jean subsequently denied making the remark, citing an interview with General Ramón Camps in which Camps states he does not believe that Saint-Jean made the comment (Camps 1983). To my knowledge, there is no public record of Saint-Jean himself denying the comment.

⁵³ *Clarín*, Dec. 18, 1977 (cited in Frontalini and Caiati 1984, p. 24).

⁵⁴ Because of the likely consequences and improbability of success of the Las Madres' actions, within Argentina they earned the reputation of "*Las Locas de la Plaza de Mayo*."

mothers, compelled them to do so (Navarro 1989, p. 256).⁴⁵ However, this explanation fails to account for the vast majority of mothers of the disappeared who did nothing to make the personal political. Most attempts to account for the actions of Las Madres do so through a NSM framework that emphasizes how "new" (previously nonpolitical),⁴⁶ "powerless" actors employ "untraditional" forms of symbolic protest and emphasize the expression of particular "identities" to make claims on society. More insightful analyses combine this approach with "an appreciation of the strategic uses of maternal legitimacy" (Brysk 1994, 187 n. 31). Additionally, the nature of repression (which "created" a new category of social actor) and the institutional context were conditioning factors (most of the women had sought help from the Church but had been turned away). In Argentina, Las Madres' appeals to representatives of the state were directed to particular government offices, where relatives of the disappeared became aware of each other and came to recognize their shared plight and the futility of seeking help through traditional channels (Brysk 1994, p. 57). In contrast, in Chile, the Agrupación de Familiares de Detenidos y Desaparecidos (AFDD) was permitted to establish its office on Church property in the Vicaría.

Las Madres are often the focus of accounts of the Argentina human rights movement; this is likely because more than any other HRO to emerge in Latin America in this period, Las Madres captured the attention and support of the international community. Whatever limited political space they occupied was facilitated by international press coverage—skillfully manipulated by Las Madres leaders (MelliBovsky 1998). The image of defenseless mothers appealing to the military regime for information on the whereabouts of their children drew support from a number of international HROs and humanitarian foundations (Brysk 1994). "Framing processes" (Snow et al. 1986) were central in linking the plight of Las Madres to the international human rights issue network (Sikkink 1993). This allowed Las Madres to access international resources and support for their struggle without the benefit of the organizational networks of the Church. The publicity campaigns of Las Madres and their sponsors in a number of international forums contributed to growing international condemnation of the junta. And though their demands were ignored by

⁴⁵ In Navarro's (1989) account, the identity of "motherhood" is the primary causal factor. This interpretation is supported by interviews with founders and members who explain their participation in such terms. However, a politicized sense of "motherhood" may be more a result of their participation in Las Madres than the original cause (see Calhoun 1991).

⁴⁶ For a critique of the claims of novelty made by many NSM theorists, see Calhoun (1993).

the junta, together with Las Abuelas,⁵⁷ these women gained increasing international publicity, influencing the process of delegitimation of the military regime, which climaxed with the Malvinas crisis in the early 1980s.

While linkages to the international human rights issue network enabled Las Madres and the broader human rights movement to survive the dictatorship and exercise limited influence in the process of transition to democracy, (Brysk 1994) their effectiveness (in terms of their own stated goals) under the military government was quite limited. In comparison with Chile, where the Church's position vis-à-vis the dictatorship and its extensive linkages to other domestic institutions and to political parties facilitated the formation of dense and extensive embedded networks of human rights activists working in a variety of spheres and levels of society, in Argentina, the HROs essentially stood alone during the worst years of repression (though collaboration among HROs facilitated limited contacts between some of the same sorts of actors who were important in the Chilean case.) In Argentina, groups of *afectados* who engaged in symbolic protest largely isolated themselves from the rest of Argentine society during the worst years of repression. In contrast, in Chile, under the auspices of the Church, religious leaders, lawyers, social workers, academics, and political party members engaged in arranging safe exile for refugees, in providing monetary, legal, and medical assistance to victims and their families, in creating soup kitchens and work programs, and in working to expand the network of NGOs, which eventually became a type of "surrogate" political opposition to the regime (Loveman 1994).⁵⁸ It becomes clear in comparative perspective that the vitality of the early HROs in Argentina was limited both by the severely repressive context and by the fact that they were not linked to previously existing social and political networks.

Conclusion

Theoretically informed comparison with Uruguay and Chile reveals Argentina to be an "intermediate case." In contrast to Uruguay, HROs

⁵⁷ Las Abuelas (grandmothers) made claims on behalf of children born to pregnant prisoners or to the disappeared. Children of the disappeared were rarely returned to their blood relatives. Instead, they were given or sold to military families or their friends.

⁵⁸ This is not meant to downplay the overall significance of the human rights movement in Argentina, both nationally and internationally; as was apparent in the transition process with its famous trials and convictions (and later pardons) of the commanding officers of the junta, the claims of Argentine HROs had a powerful impact on political culture (even if their attempts to influence policy largely failed; see Brysk 1994).

emerged in Argentina both prior to and during the height of repression. However, the Argentine organizations lacked the foundation of extensive social networks, connected to domestic and international institutional networks of the Catholic Church, which characterized Chilean HROs. HROs in Argentina were thus much more vulnerable and hence less effective in terms of their own goals than their counterparts in Chile. Personal networks linking non-Catholic and dissident Catholic religious leaders to each other and to their parishes, and the previously developed prosocial religious and moral commitments of these leaders, enabled the development of SERPAJ and MEDH in response to the repression prior to the coup (1974–76). Individual politicians, lawyers, professionals, and religious leaders created the APDH as a (noncommunist) alternative civil-libertarian HRO (Mignone 1991, pp. 99–106). These groups, as well as *Las Madres* and *Las Abuelas*, emerged despite (or because of) a severely unfavorable political opportunity structure by operating within “noninstitutional” spaces that were not *formally* restricted. However, these spaces offered little or no protection from state persecution, restricting the development and activities of HROs under the dictatorship. Comparison with the extensive programs under COPACHI and later the *Vicaría* in Chile suggests the crucial significance of such a protected space for the effectiveness and sustainability of contentious collective action in high-risk contexts. The absence of Church support also limited the possibilities for the HROs to develop networks linking activists from different sectors of society; this further restricted the effectiveness of these organizations in terms of their own goals. This was particularly evident for *Las Madres* and *Las Abuelas*. Despite international fame, the domestic influence of these groups during the mid- to late 1970s was limited by their relative isolation from other sectors of society.⁹⁹

CONCLUSION

By asking why and under what conditions individuals will risk their lives to confront state repression, I have identified an important area of research largely neglected in the existing social movement literature. Contrary to sociological and commonsense expectations, the cases examined

⁹⁹ The relative lack of domestic networks linking Argentine HROs to other social sectors, such as political parties, universities, and unions, may also partly account for their lesser influence during the transition to democracy, as compared to Chile. Comparative analysis of the role of HROs in the transitions in these countries, as well as their fate in newly democratic contexts, merits future research. For insights on Chile and Uruguay, see de Brito (1997); on Chile and Argentina, see Skaar (1994); on Chile in Comparative Perspective see Frühling and Orellana (1991).

here demonstrate that the onset of severe state repression, that increases dramatically both the potential risks and costs of collective action, may itself stimulate certain types of social movements. The generalized demobilization that is the expected outcome of dramatic increases in the scope and scale of state repression does not capture the entire picture; state repression may stimulate collective organization and opposition from certain sectors of society as a direct result of the severity and cruelty of its attempts to stifle it in others.

Comparative analysis suggests that participation in high-risk collective action depends largely on particular types of personal social network ties (McAdam 1986) and the particular way in which face-to-face networks are embedded within broader institutional networks (Morris 1984; della Porta 1988). Reliance on face-to-face networks permits a high degree of trust that helps to counteract the selective *disincentives* to participate posed by threats of state persecution. Dense interpersonal networks tend to insulate activists, which contributes to their intensified commitment and willingness to act despite risks of horrific repercussions. The particular way in which certain types of face-to-face networks are embedded within previously existing domestic and international networks largely determines the ability of activists to organize and sustain collective action in high-risk contexts. To explain the actions of the “early risers” in high-risk contexts, it is also important to consider how forms of repression collide with personal, moral, or political commitments developed prior to and in the process of participation (Calhoun 1991; Martín-Baró 1983). The diversity and density of networks within which the personal networks of early risers are embedded influences whether, or to what extent, the efforts of the first core group of actors will develop into sustained and effective collective resistance and opposition, creating openings for later waves of activists by reducing the costs or risks associated with joining. But ultimately the emergence and sustainability of social movement organizations is constrained by the political opportunity structure, particularly the levels and types of repression—that is, the ability of the government to effectively curtail and control access to material and symbolic resources and physical and sociopolitical space.

When early risers are able to create or expand openings, whether they will be filled by new actors depends in part on the ability to frame the struggle in terms that resonate in the wider society. Framing processes are also important for linking local struggles into broader “international issue networks” that may generate international publicity and funds (Snow et al. 1986; Snow and Benford 1988; Sikkink 1993). Additionally, both timing and participation are intimately linked to how the structure of nested social networks creates or impedes access to crucial resources including physical, sociopolitical, and symbolic ‘space,’ information, and

material assistance, particularly from transnational and international sources.

Comparative historical analysis employing a synthetic theoretical framework facilitates identification of those social processes that influence the emergence of collective action in high-risk contexts that are generalizable across cases with the consequent potential to *generate theory*. This approach illuminates how *particular configurations* of 'variables' affect outcomes. In an example from this article, comparison of Argentina and Chile informed by a synthetic theoretical approach suggested that the emergence and effectiveness of HROs was related not merely to the presence or absence of particular types of personal ties, *plus* the presence or absence of institutional networks, but rather to the particular way in which personal networks were *embedded within* broader, multilevel institutional networks. This approach redirects attention of the theorist to *relationships among* social processes operating at different levels, resulting in improved understanding of the linkages between the how and the why of social movements.

This approach thus offers the potential to bridge some of the gaps between resource mobilization and NSM theoretical schools, a need that has been increasingly recognized and articulated by leading theorists in the field (Klandermans and Tarrow 1988; McAdam, Tarrow, and Tilly 1995). It also facilitates borrowing from, and integrating, the contributions of earlier pluralist, rational choice, structuralist, marxist, and social-psychological social movement theorists. If construction of social theory and social scientific knowledge is to be a cumulative enterprise, the contributions from diverse theories and comparative historical research must be incorporated and assimilated rather than discarded or forgotten in a battle of competing paradigms. The explanations generated through this comparative historical analysis of the emergence of HROs in Chile, Uruguay, and Argentina thus serve as a modest example of the utility of synthetic theory building and comparative research rather than pseudoparadigmatic intellectual warfare.

TABLE A1
ESTIMATED EXTENT OF REPRESSION IN THREE SOUTHERN CONE COUNTRIES

COUNTRY	TIME PERIOD	PRISONERS									
		DEATHS			LONG-TERM			CUMULATIVE			EXILES
		Total	Per 1,000 Population		Total	Per 1,000 Population		Total	Per 1,000 Population		
Argentina	1976-83	10,000	40		7,000	03		30,000	11		26,480,000
Chile ..	1973-77	4,000	.40		6,500	.60		60,000	5.8		10,372,000
Uruguay	1973-84	36	.01		4,000	1.40		60,000	21.0		2,847,000

SOURCE: Repression statistics are from King (1989); population statistics are from United Nations (1987)

NOTE: King's estimates are derived from a comparative analysis of a wide range of estimates from government sources, international and domestic human rights organization sources, and academic sources.

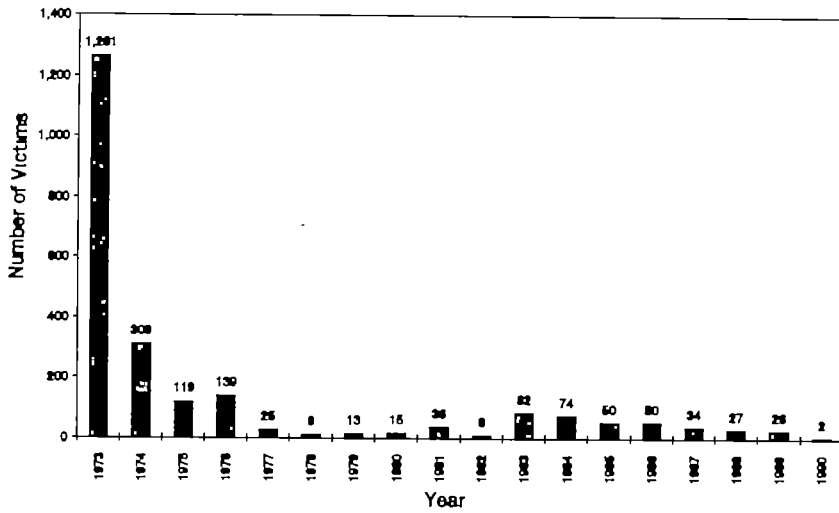


FIG. A1.—Victims of death and disappearance in Chile, 1973–90 (Chilean National Commission on Truth and Reconciliation 1993, p. 903).

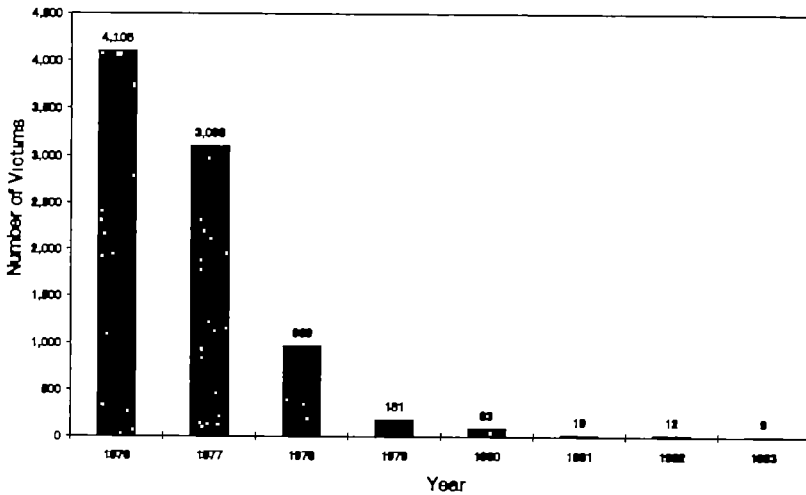


FIG. A2.—Victims of disappearance in Argentina, 1976–83 (Sikkink 1993, p. 427; Skaar 1994).

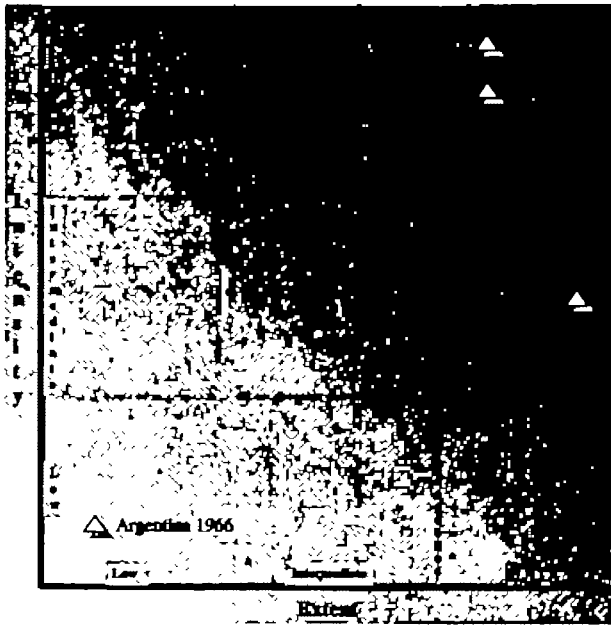


FIG. A3.—Comparative intensity and extent of repression (from King 1989); "extent" refers to the gross number of individuals affected by repression; "intensity" refers to the manner in which they were affected. Extent is the cumulative total of deaths, number of long-term political prisoners, total number of political prisoners, and number of political exiles. Intensity is determined by the number of deaths and long-term imprisoned, with greater weight given to deaths. Torture is not systematically incorporated into King's analysis (it is notoriously difficult to "measure"). However, he does mention that "qualitative aspects of the different cases are also evaluated in making the rankings."

REFERENCES

- Amnesty International. 1983. "Mental Health Aspects of Political Imprisonment in Uruguay: An Amnesty International Special Briefing." New York: Amnesty International.
- Bravo Lira, Bernardino. 1978. *Régimen de gobierno y partidos políticos en Chile, 1924–1973*. Santiago: Editorial Jurídica de Chile.
- Brysk, Alison. 1994. *The Politics of Human Rights in Argentina: Protest, Change, and Democratization*. Stanford, Calif.: Stanford University Press.
- Calhoun, Craig. 1991. "The Problem of Identity in Collective Action." Pp. 51–75 in *Macro-Micro Linkages in Sociology*, edited by Joan Huber. Newbury Park, Calif.: Sage.
- . 1993. "New Social Movements of the Early Nineteenth Century." *Social Science History* 17: 385–427.
- Camps, Ramón J. 1983. "Los desaparecidos están muertos." *El Bimestre Político y Económico* 7 (January–February): 62–64.
- Camus L., Carlos. 1985. "La experiencia de la Iglesia Chilena en la defensa de los

- derechos humanos." Pp. 51–58 in *Represión política y defensa de los derechos humanos*, edited by Hugo Frühling. Chile: Academia de Humanismo Cristiano.
- Centro de Estudios Legales y Sociales. 1984. "Uruguay/Argentina: Coordinación represiva." Buenos Aires.
- CEP (Centro de Estudios y Publicaciones) 1974. *Chile: Masacre de un pueblo*. Lima: Centro de Estudios y Publicaciones
- Chilean National Commission on Truth and Reconciliation. 1993. *Report of the Chilean National Commission on Truth and Reconciliation*, vol. 2. Translated by Phillip E. Berryman, with an introduction by Jose Zalaquett. Notre Dame, Ind.: Notre Dame University Press.
- Comisión Argentina por los Derechos Humanos. 1977. *Argentina: Proceso al genocidio*. Madrid: Elías Querejeta Ediciones.
- Corradi, Juan E., Patricia W. Fagen, and Manuel A. Garretón, eds. 1992. *Fear at the Edge: State Terror and Resistance in Latin America*. Berkeley and Los Angeles: University of California Press.
- de Brito, Alexandra B. 1997. *Human Rights and Democratization in Latin America: Uruguay and Chile*. New York: Oxford University Press.
- della Porta, Donatella. 1988. "Recruitment Processes in Clandestine Political Organizations: Italian Left-Wing Terrorism." Pp. 155–69 in *International Social Movement Research*. Vol. 1. Greenwich, Conn.: JAI Press.
- Diani, Mario. 1996. "Linking Mobilization Frames and Political Opportunities: Insights from Regional Populism in Italy." *American Sociological Review* 61 (6): 1052–69.
- Drake, Paul W. 1996. *Labor Movements and Dictatorships: The Southern Cone in Comparative Perspective*. Baltimore: Johns Hopkins University Press.
- Duhade, Eduardo Luis. 1983. *El estado terrorista argentino*. Spain: Talleres Gráficos.
- Fireman, Bruce, and William Gamson. 1979. "Utilitarian Logic in the Resource Mobilization Perspective." Pp. 8–44 in *The Dynamics of Social Movements: Resource Mobilization, Social Control, and Tactics*, edited by Mayer N. Zald and John D. McCarthy. Cambridge, Mass.: Winthrop Publishers
- Fitzgibbon, Russell. 1954. *Uruguay: Portrait of a Democracy*. New Brunswick, N.J.: Rutgers University Press.
- Franco, Rolando. 1984. *Democracia "a la Uruguaya": Análisis electoral, 1925–1985*. Montevideo: Editorial El Libro Libre.
- Frontalini, Daniel, and María C. Calati. 1984. *El mito de la "Guerra Sucia"*. Buenos Aires, Argentina: Centro de Estudios Legales y Sociales.
- Frühling, Hugo. 1988. "Organismos no gubernamentales de derechos humanos en el paso del autoritarismo a la democracia en Chile." Pp. 137–65 in *Una puerta que se abre*, Taller de Cooperación al Desarrollo. Chile: Imprenta la Unión.
- Frühling, Hugo, Gloria Alberti, and Felipe Portales. 1989. *Organizaciones de Derechos Humanos de América del Sur*. San José, Costa Rica: Instituto Interamericano de Derechos Humanos.
- Frühling, Hugo, and Patricio Orellana. 1991. "Organizaciones no gubernamentales de derechos humanos bajo regímenes autoritarios y en la transición democrática: El caso chileno desde una perspectiva comparada." Pp. 25–83 in *Derechos Humanos y Democracia*, edited by Hugo Frühling. Chile: Instituto Interamericano de Derechos Humanos.
- García, Prudencio. 1995. *El drama de la autonomía militar*. Madrid: Alianza Editorial.
- Garretón, Manuel A. 1983. *El proceso político Chileno*. Santiago: Talleres Gráficos.
- Gauding, Anna-Karin. 1991. *Es mejor encender una luz que maldecir la oscuridad: Sobre el trabajo de Diakonia por los derechos humanos en América Latina*. Santiago: Diakonia
- Geddes, Barbara. 1995. "Uses and Limitations of Rational Choice." Pp. 81–108 in *Latin America in Comparative Perspective: New Approaches to Methods and Analysis*, edited by Peter Smith. Boulder, Colo.: Westview Press.

- Gould, Roger V. 1995. *Insurgent Identities: Class, Community, and Protest in Paris from 1848 to the Commune*. Chicago: University of Chicago Press.
- Gurr, Ted R. 1970. *Why Men Rebel*. Princeton, N.J.: Princeton University Press.
- Hirsch, Eric L. 1986. "The Creation of Political Solidarity in Social Movement Organizations." *Sociological Quarterly* 27 (3): 373-87.
- Inter-Church Committee on Human Rights in Latin America. 1978. *Violations of Human Rights in Uruguay (1972-1976)*. Toronto: Inter-Church Committee on Human Rights in Latin America.
- Kaufman, Edy. 1979. *Uruguay in Transition: From Civilian to Military Rule*. New Brunswick, N.J.: Transaction Books.
- King, Peter J. 1989. "Comparative Analysis of Human Rights Violations under Military Rule in Argentina, Brazil, Chile, and Uruguay." Pp. 1043-65 in *Statistical Abstract of Latin America*, vol. 27. Edited by James W. Wilkie and Enrique Ochoa. Los Angeles: University of California, UCLA Latin American Center Publications.
- Klandermans, Bert. 1984. "Mobilization and Participation: Social-Psychological Expansions of Resource Mobilization Theory." *American Sociological Review* 49 (5): 583-600.
- Klandermans, Bert, and Sidney Tarrow. 1988. "Mobilization into Social Movements: Synthesizing European and American Approaches." Pp. 1-38 in *International Social Movement Research*, vol. 1. Greenwich, Conn.: JAI Press.
- Laitin, David D. 1995. "National Revivals and Violence." *Archives Europeennes de Sociologie* 36:3-43.
- Lawyers Committee for International Human Rights. 1985. "The Generals Give Back Uruguay: A Report on Human Rights." New York: Lawyers Committee for International Human Rights.
- Lindahl, Goran G. 1964. "Uruguay: Government by Institutions." Pp. 447-61 in *Political Systems of Latin America*, edited by Martin C. Needler. Princeton, N.J.: D. Van Nostrand.
- Lipset, Seymour M., and Earl Raab. 1970. *The Politics of Unreason: Right Wing Extremism in America, 1790-1970*. New York: Harper & Row.
- Lira, Elizabeth. 1997. E-mail interview with author February 15.
- Lofland, John. 1996. *Social Movement Organizations: Guide to Research on Insurgent Realities*. New York: Aldine de Gruyter.
- Loveman, Brian. 1994. "Las ONG Chilenas: Su papel en la transición a la democracia." *Nuevas Políticas Urbanas*, edited by Charles A. Reilly. Arlington, Va.: Fundación Interamericana.
- Lowden, Pamela. 1996. *Moral Opposition to Authoritarian Rule in Chile, 1973-90*. New York: Macmillan.
- Mainwaring, Scott. 1986. *The Catholic Church and Politics in Brazil, 1916-1985*. Stanford, Calif.: Stanford University Press.
- Martín-Baró, Ignacio. 1983. *Acción e ideología psicología social desde Centroamérica*. San Salvador, El Salvador: Talleres Gráficos UCA.
- Marwell, Gerald, and Pamela Oliver. 1984. "Collective Action Theory and Social Movements Research." *Research in Social Movements, Conflict and Change* 7:1-27.
- McAdam, Doug. 1986. "Recruitment to High-Risk Activism: The Case of Freedom Summer." *American Journal of Sociology* 92 (1): 64-90.
- McAdam, Doug, Sidney Tarrow, and Charles Tilly. 1995. "To Map Contentious Politics." Working Paper no. 222. New School for Social Research, Center for Studies of Social Change.
- MelliBovsky, Matilde. 1998. Circle of Love over Death. Lecture given at the Latin American Studies Student Organization Conference on Human Rights in the Americas. April 14, at San Diego State University, San Diego, California.

- Melucci, Alberto. 1988. "Getting Involved: Identity and Mobilization in Social Movements." Pp. 329-48 in *International Social Movement Research*, vol. 1. Greenwich, Conn.: JAI Press.
- Mignone, Emilio F. (1986) 1988. *Witness to the Truth: The Complicity of Church and Dictatorship in Argentina, 1976-1983*, translated by Phillip Berryman. Maryknoll, N.Y.: Orbis Books.
- . 1991. *Derechos humanos y sociedad: El caso Argentino*. Buenos Aires: Centro de Estudios Legales y Sociales; Ediciones del Pensamiento Nacional.
- Morris, Aldon. 1984. *The Origins of the Civil Rights Movement: Black Communities Organising for Change*. New York: Free Press.
- Navarro, Marysa. 1989. "The Personal Is Political: Las Madres de la Plaza de Mayo." Pp. 241-58 in *Power and Popular Protest: Latin American Social Movements*, edited by Susan Eckstein. Berkeley and Los Angeles: University of California Press.
- Oberschall, Anthony. 1973. *Social Conflict and Social Movements*. Englewood Cliffs, N.J.: Prentice-Hall.
- . 1993. *Social Movements: Ideologies, Interests and Identities*. New Brunswick, N.J.: Transaction.
- Ollver, Pamela. 1984. "If You Don't Do It, Nobody Else Will": Active and Token Contributors to Local Collective Action." *American Sociological Review* 49 (5): 601-10.
- Olson, Mancur. 1965. *The Logic of Collective Action: Public Goods and the Theory of Groups*. Cambridge, Mass.: Harvard University Press.
- Orellana, Patricio, and Elizabeth Q. Hutchison. 1991. *El movimiento de derechos humanos en Chile, 1973-1990*. Santiago: Centro de Estudios Políticos Latinoamericanos Simón Bolívar.
- Osa, Maryjane. 1989. "Resistance, Persistence, and Change: The Transformation of the Catholic Church in Poland." *Eastern European Politics and Societies* 3 (2): 268-99.
- Rama, German W. 1987. *La democracia en Uruguay*. Buenos Aires: Grupo Editor Latinoamericano.
- Rock, David. 1985. *Argentina, 1516-1987*. Berkeley and Los Angeles: University of California Press.
- Servicio Paz y Justicia. 1989. *Uruguay nunca más: Human Rights Violations, 1972-1985*, translated by Elizabeth Hampsten. Philadelphia: Temple University Press.
- Sikkink, Kathryn. 1993. "Human Rights, Principled Issue-Networks, and Sovereignty in Latin America." *International Organization* 47 (3): 411-41.
- Skaar, Elin. 1994. "Human Rights Violations and the Paradox of Democratic Transition." Master's thesis. University of Bergen, Department of Comparative Politics.
- Smith, Brian H. 1982. *The Church and Politics in Chile. Challenges to Modern Catholicism*. Princeton, N.J.: Princeton University Press.
- Snow, David, and Robert Benford. 1988. "Ideology, Frame Resonance, and Participant Mobilization." Pp. 197-217 in *International Social Movement Research*. Vol. 1. Greenwich, Conn.: JAI Press.
- . 1992. "Master Frames and Cycles of Protest." Pp. 133-55 in *Frontiers in Social Movement Theory*, edited by Aldon Morris and Carol M. Mueller. New Haven, Conn.: Yale University Press.
- Snow, David, E. Burke Rochford, Jr., Steven Worden, and Robert Benford. 1986. "Frame Alignment Processes, Micromobilization, and Movement Participation." *American Sociological Review* 51 (4): 464-81.
- Tarrow, Sidney G. 1994. *Power in Movement: Social Movements, Collective Action, and Politics*. Cambridge: Cambridge University Press.
- Tilly, Charles. 1978. *From Mobilization to Revolution*. Reading, Mass.: Addison-Wesley.

- United Nations. 1987. *1985 Demographic Yearbook*, 37th issue. Department of International Economic and Social Affairs, Statistical Office. New York: United Nations.
- Veiga, Raúl. 1985 *Las Organizaciones de derechos humanos*. Buenos Aires: Centro Editor de America Latina.
- Vidal, Hernan. 1986. *El Movimiento contra la tortura "Sebastián Acevedo": Derechos humanos y la producción de símbolos nacionales bajo el fascismo Chileno*. Edina, Minn.: Institute for the Study of Ideologies and Literature, Society for the Study of Contemporary Hispanic and Lusophone Revolutionary Literatures.
- Weber, Max. (1922) 1968. *Economy and Society*. Berkeley: University of California Press.
- Weschler, Lawrence. 1990 *A Miracle, a Universe: Settling Accounts with Torturers*. New York: Pantheon Books.
- Wickham-Crowley, Timothy P. 1989. "Winners, Losers, and Also-Rans Toward a Comparative Sociology of Latin American Guerrilla Movements." Pp. 132-81 in *Power and Popular Protest: Latin American Social Movements*, edited by Susan Eckstein. Berkeley and Los Angeles: University of California Press.
- Zald, Mayer N., and John D. McCarthy, eds. 1979. *The Dynamics of Social Movements: Resource Mobilization, Social Control, and Tactics*. Frontiers of Sociology Symposium. Cambridge, Mass. Winthrop.

Book Reviews

Of Revelation and Revolution. Vol. 2, *The Dialectics of Modernity on a South African Frontier*. By John L. Comaroff and Jean Comaroff. Chicago: University of Chicago Press, 1997. Pp. xxiii+560. \$70.00 (cloth); \$24.95 (paper).

David Chidester
University of Cape Town

In this second installment of a three-volume project, John L. Comaroff and Jean Comaroff advance their historical ethnography of the colonial encounter between Southern Tswana and the 19th-century British "civilizing mission" along the contested frontiers of southern Africa and social theory. While the first volume that appeared in 1991 usefully developed and insightfully deployed its terms of analysis—culture, ideology, and hegemony—in order to reenter the colonization of consciousness and the consciousness of colonization, the second volume more directly addresses colonialism as a nexus for the production and performance of everyday practices. Simultaneously material and cultural, the everyday activities entailed in cultivation and exchange, clothing and housing, healing and hygiene, law and custom reveal a colonial "epic of the ordinary" in which subjectivities were renegotiated in ways that had distinctively "modernizing" effects at both the colonized periphery and the European metropolitan center. By turning on this dialectic of the ordinary and the extraordinary, which Comaroff and Comaroff gloss as the "epiphanies of the everyday," the book in effect develops a sociology of religion that takes as its implicit theoretical point of departure a Marxist characterization of capitalist commodification as the modern "religion of everyday life." Obviously, the British colonial missions carried into southern Africa an extraordinary religion of sin and salvation, a universalizing religion (using "religion," as the authors note, "in the restricted sense of the term") that on its own criteria of success—widespread conversion—was a failure. Dramatic changes in the religion of everyday life, however, were produced, in the phrase used by the 1820s London Missionary Society agent John Philip, through a "revolution in habits." Although they position their investigations between the European age of revolution (1789–1848) and the negotiated revolution of the 1990s in South Africa, the authors are really interested in this volume in reporting from the frontlines of the ordinary revolution of habits in modernity.

While reviewing the extraordinary religion of the Christian mission in chapter 2, the authors devote chapters 3–8 to a detailed and evocative analysis of that ordinary revolution of everyday life. In and through the

Permission to reprint a book review printed in this section may be obtained only from the reviewer.

close readings contained in these chapters on agricultural production, money, clothing, architecture, healing, and the human agency modulated by the contradictory ideological constructions of the modernist citizen and the ethnic subject, the authors develop their dialectical arguments about colonialism. Not merely a function of political economy, colonialism is also a cultural formation, a discursive field, or a regime of knowledge, strategies, and techniques that arises out of the "mutually transforming play of social forces." If the analysis rests more heavily on discourses than on forces, if it immerses itself in conversations, contradictions, and convergences rather than within the brute social physics of coercion, within the colonialism that Ashis Nandy has identified as the "armed version" of modernism, then it should nevertheless be appreciated as a necessary corrective to more mechanical renderings of colonial situations. This volume, like its predecessor, is a landmark contribution to social history and social anthropology that is neither, yet also both, at the same time. It should be read by anyone interested not merely in colonialism but also in the social.

Still, critics complain. One of the delights of this book is reading the brief opportunity that the authors take to respond to critics of the first volume, even though they graciously invite readers to skip over these pages. Although many reviewers were supportive and collaborative, some were argumentative and difficult. While I am generally supportive of this second volume, I raise three problems. First, the problem of power, which the Comaroffs feel they can resolve through a dialectic of mutual determinations in colonial situations, will seem unresolved for readers who prefer to emphasize either the social physics of colonization or the historical agency of the colonized. Although the Comaroffs carefully restate their dialectical position, this problem of power still seems at an impasse. Second, I find that the Comaroffs are well positioned to advance a theory not only of colonialism but also of religion. While religion is used as a rhetorical trope, as in the religion of the everyday, religion is not explicitly problematized as a category emerging from the very colonial encounters that the authors have so usefully documented. Finally, having learned so much from the first volume in 1991, I have to admit that I was not surprised by what I read in the second volume. Why do I want to be surprised? It is not just because of my modernist subjectivity, which has arisen in part out of a commodification of novelty, but it is also a result of the expectations raised by the exhilarating attention to theory that the Comaroffs exemplified in their first volume. Theory, it has been said, is an instrument of surprise: no surprise, no theory—no matter how theoretically sophisticated. I could not help feeling in reading this second volume that we were moving through familiar territory rather than new frontiers. Perhaps that sense of *déjà vu* reflects the extent to which I along with many others have integrated the authors' historical ethnography into my understanding of "normal" scholarship. Nevertheless, while the first volume registered as a powerful and strategic intervention in social theory, the second seems more like a process of consolidation. Looking forward

to their promised third volume on colonial pedagogy, however, I still nurture the hope of being once again surprised.

Struggle for the Spirit: Religious Transformation and Popular Culture in Brazil and Latin America. By David Lehmann. Cambridge: Polity Press, 1996. Pp. xiv+244. \$45.00.

John Burdick
Syracuse University

In the 1980s, by his own account, David Lehmann's researches in Latin America had led him to a strong interest in the progressive Catholic Church. Moved by the example of Nicaragua, he found himself seeking to understand the meaning and resonances of the church's message to the "people" throughout the continent. By the 1990s, however, he had become increasingly aware of the other major spiritual force in Latin America: the Protestants. What had begun as an enquiry among the Catholics now expanded. What were the relationships, he wanted to know, between these two religious movements and "people's culture" in Latin America? "How does their language," he asked, "relate to that of the people? How do their metaphors, symbols and rituals relate to the traditions of the people's culture? How do they themselves 'think the people'? And how are these radically opposed projects of salvation . . . interpreted in the language of the people?" (p. 4).

Posing questions in this open-ended way has led Lehmann to make, in this engaging book, a significant contribution to the rapidly growing literature on the "Catholic-Protestant" question in Latin America. In particular, this open-endedness permitted Lehmann to avoid seeing the two movements exclusively in relation to each other as "winners" or "losers" in the battle for souls. Instead, drawing on his research in the Brazilian northeast, and on his knowledge of other parts of Latin America, Lehmann has provided us with a carefully detailed account of the ideas and practices espoused by the leadership cadres of the two movements, as each engages, in its own way, in the effort to speak to the masses. The result is a nuanced portrait of activists in the two movements in the process of articulating ideologies, building institutions, and mobilizing resources.

Lehmann begins with an important discussion of what he appropriately calls the discourse of Catholic *basismo*, the romantic, committed engagement with "the people." He sees this discourse as belonging to a long Catholic tradition, which includes other movements such as Opus Dei and Catholic Action, of striving to connect with and mobilize mass constituencies. At the same time, he argues that *basismo* has developed a practical emphasis on the notion of the quality as opposed to the quantity of followers. In this context the chapters on the intricacies of *basista* language and on the everyday practices of raising the consciousness of

the laity break new ground. In contrast to many of the critiques of the progressive Catholic Church that portray it as overly intellectual and abstract, Lehmann is able to show that *basismo* has led to the formation of a nonelite intelligentsia able and willing to accommodate a range of popular beliefs and practices—such as saint worship—that only a few years ago were anathema to the radical Catholic Church. In the final analysis, however, Lehmann sees *basismo* as caught in the contradiction of glorifying the popular while retaining ultimate authority in the hands of intellectuals, elite or not. This contradiction leads, in Lehmann's felicitous phrase, to an ongoing sense of "unease," of "anxiety to get ever nearer the pits of poverty, and a nagging sense that those pits are unreachable" (p. 98).

Lehmann's discussion of pentecostals is equally complex. He points out that pentecostalism's relationship to "the people" is ambivalent. On the one hand, it derives much of its strength from a systematic rejection and transcendence of popular culture, by distancing itself not only from the forms of popular Catholicism, but also from the whole range of masculinist norms that disempower women (p. 194). On the other hand, pentecostal churches make use of a variety of familiar popular practices, including those derived from the possession cults (p. 137).

In the end, however, Lehmann regards pentecostalism as possessing greater potential than *basista* Catholicism as a force for cultural change because, while the Catholics continue to reinforce the authority of intellectuals as articulators of the popular, the pentecostals have taken an (unintended) historical step: "They reject the image of the people as projected by the intelligentsia, but they also reject the intelligentsia with great force." Consequently, "they constitute a force as potent as it is disconcerting" (p. 228).

It is fitting that Lehmann would end a book dedicated to challenging both the romance of *basismo* and liberal anxiety about the pentecostals on this politically ambiguous note. Yet it is also appropriate to point out that Lehmann's work itself participates more in an intellectual rather than an anti-intellectual tradition of social criticism. The lion's share of Lehmann's sources are, after all, the theologians, priests, pastors, and "organic intellectuals" who have taken it upon themselves to formulate and articulate (to, among others, intrepid researchers) the origins, meaning, and destinies of the religious movements they lead. Still marginal here, despite Lehmann's good intentions, are the voices of "the people" who, at the start, we were alerted would be consulted this time around. It is quite possible that Lehmann did in fact consult with many of them. However, by not working into his text what most probably was the complex, contradictory range of views he heard from them, Lehmann may have indirectly contributed to a critique of intellectualism in social movements: by speaking once again "about" nonintellectuals, he may have strengthened our will, and theirs, to hear nonintellectuals speak for themselves.

Disquiet in the Land: Cultural Conflict in American Mennonite Communities. By Fred Kniss. New Brunswick, N.J.: Rutgers University Press, 1997. Pp. xiii+257. \$50.00 (cloth); \$19.95 (paper).

Courtney Bender
Princeton University

Fred Kniss's sophisticated new book about conflicts in Mennonite communities does much more than chronicle the irony of the self-proclaimed "quiet in the land" incessantly and vociferously arguing over cultural symbols. It also advances our understanding of the interplay between organizational and cultural resources, demonstrating how, over time, signs and symbols are both objects of conflict and resources that are mobilized by contending parties. Kniss uses a historical narrative and "qualitative comparative analysis" (based on Boolean algebra) to describe and assess the population of conflict events in American Mennonite communities between 1870 and 1985. In doing so, he also considers several issues of general interest for cultural sociologists.

Kniss situates his study on two premises: cultural symbols, signs, and ideologies are not epiphenomenal, and religious groups are not closed systems. At a basic level, these positions address gaps in existing literature on intrareligious conflict, but they also guide Kniss to questions that are broader in scope. For instance, how do cultural signs and symbols play a role in structuring conflict events? Furthermore, how are larger historical events and larger ideological systems important in local settings?

Analyzing events and locating each within particular social and historical contexts allows Kniss to discuss some of the questions about the interplay between Mennonite conflicts and larger social trends. The first part of the book describes four distinctive periods of conflict over a 115-year period. While conflicts do not mirror the ideological and symbolic debates that go on in the larger society, they are nevertheless patterned in ways that demarcate ongoing interaction between large-scale social change and the internal dynamics of Mennonite communities. Mennonite leaders sometimes drew strategically on external ideological sources and symbols, while at other times larger social events limited the degrees to which Mennonites looked outward for symbolic resources. Over time, the changing location of legitimate authority, the power of particular symbols within the larger domain of Mennonite ideological debates, and the increasing complexity of Mennonite organizational forms all influence the shape of emerging conflict events.

While these forces intersect to help structure conflict events, they are nevertheless not equal partners at all stages of a conflict. Cultural resources, embedded in networks of external and internal social forces, clearly influence the *emergence* of conflict events. Challengers and defenders choose strategies based on the available cultural resources and the ways they are valued within larger sociocultural environments. (Here, Kniss develops a much needed multidimensional map of mainstream

American ideological terrain, moving far beyond the normal bipolar models often used in analysis of contemporary "culture wars." This model is in itself a worthy contribution to current debates in the sociology of culture and religion.) Yet once a conflict is set in motion, organizational factors including the degree to which defenders are organized and the presence of third parties become more important than cultural ones in determining its outcome. Kniss finds that cultural content matters little in determining the probable *outcome* of conflicts (whether conflicts will wither away, or end in victory, schism, defeat, compromise).

Kniss is surprised by this finding and notes that his method favors organizational over cultural variables. Kniss's qualitative comparative method allows him to use his full set of conflict events in identifying the factors involved in different conflict outcomes. It nevertheless also requires him to code culture as ideological content of conflicts. Coding culture in such a way glosses over many insights about the interaction of cultural resources and organizational forms that Kniss makes in the narrative portion of the text. For instance, he describes how the role of third-party interventions changed through time. We learn how third-party membership changed as Mennonite denominational structures grew, and we hear how their strategies changed from adjudicatory to conciliatory. This suggests that the symbolic value of third-party intervention changed as well, reinforcing the notion that the effects of third-party interventions are not "merely" organizational. In Kniss's descriptions of these changes, we see how organizational forms are inherently ideologically loaded. I imagine that cultural resources embedded in organizational forms play a larger role in conflict outcomes than Kniss's analysis suggests. As Kniss himself notes, his archival sources do not offer up the kind of data that further microlevel analysis would require.

The fact that Kniss's analysis provokes such questions is testament to its larger promise for several sociological fields. It prods sociologists of religion to consider how multidimensional models of religious conflict and legitimacy can broaden understandings of religious change. It also demonstrates how analyzing the changing salience of particular cultural resources can deepen our understanding of the emergence of conflicts and social movements. Most promisingly, Kniss presents a new case into the study of social and cultural change. This meticulous empirical analysis demonstrates what we have to gain by looking closely at the dynamic interactions between culture and structure that unfold through time.

Reinventing American Protestantism: Christianity in the New Millennium. By Donald E. Miller. Berkeley and Los Angeles: University of California Press, 1997. Pp. ix+253. \$27.50.

Nancy L. Eiesland
Emory University

Rational choice theorists, including those who designate themselves "new paradigm" thinkers, have attributed the growth and decline of religious organizations to the machinations of a free market religious economy, characterized by pluralism and competition. In *Reinventing American Protestantism*, Donald Miller draws extensively on this school of thought to account for the growth of an emergent group of religious organizations. Designating them "new paradigm" churches, Miller concentrates on three West Coast-based groups—Calvary Chapel, the Vineyard Christian Fellowship, and, to a lesser extent, Hope Chapel.

These upstart "postmodern sects"—Miller's term—have a simple and intensely spiritual message that makes high demands on members. These rapidly growing groups place a strong emphasis on biblical authority, a conversionist approach to evangelism, and culturally current and emotionally affecting worship. Throughout the book, these "winners" in the spiritual marketplace are contrasted with the "losers," that is, mainline Protestant denominations whose theological refinement and lack of cultural currency keep them from attracting their share of the baby boom and generation X age groups. This comparison yields both provocative proposals for the mainline denominations, delineated in the concluding chapter, and some glaring stereotypes of the losers.

The strength of *Reinventing American Protestantism* is its sympathetic portrayal of the new paradigm churches. Miller is at his best when describing his response, as a progressive Episcopalian, to the services and people of these new evangelical movements. Without self-conscious postmodern ado about the author's place in the text, Miller integrates the personal and the ethnographic narrative in an engaging manner. In one particularly moving account, Miller depicts his experience at a Vineyard Christian Fellowship healing conference. He interprets the actions and emotions of the scene—which included energetic music making and straightforward testimonies by an Indian healer/evangelist, interspersed with much weeping, laughing, and being "slain in the spirit"—while also revealing his own anxieties and conflicted feelings.

Along with fellow researchers Brenda Brasher and Paul Kennedy, Miller gathered detailed qualitative and quantitative data from 200 interviews, extensive participant observation, and surveys with more than 3,000 congregants and 400 pastors. Though missing information about the sampling procedures for the interviews and surveys and omitted frequency distributions on variables frustrates the reader, these lapses are mostly compensated for by the richness of the ethnographic detail.

The following picture of these new paradigm churches emerges.

Founded and guided by largely self-taught male ministers, these congregations attract predominately white middle-class individuals who have some college education, most of whom work in service sector jobs. Most church members are married couples with at least one child. They tend to hold conservative political views, especially on contentious issues such as abortion, homosexuality, and sex education. Many have experienced biographical disruptions, such as divorce and addiction. Miller found that 30% were reared in homes where there had been a divorce; the same number reported that their parents abused alcohol or drugs (p. 109).

Despite the similarities among the adherents to these movements, Miller highlights differences among the groups as well. Calvary Chapel draws primarily a working- and lower-middle-class constituency with a preference for conservative theology, while Vineyard Christian Fellowship's somewhat more affluent and better-educated adherents are attracted to more moderate and expressive charismatic/pentecostal theology. Occupying a middle ground theologically between Calvary Chapel and Vineyard, Hope Chapel has been particularly successful in Hawaii where it draws the most ethnically diverse crowd of the three.

The organizational patterns of these groups also diverge. Calvary Chapel has sought to maintain "a loose fellowship of like-minded people" (p. 34) with founder Chuck Smith exercising considerable informal influence but little structural authority over the more than 600 affiliated churches in the United States. Vineyard, which numbers more than 400 affiliated stateside congregations, identifies itself as a denomination with a more explicit authority structure. Founder John Wimber sought to keep ties with affiliated pastors relational rather than formal; however, Wimber's untimely death in 1997 may create leadership challenges for the young denomination. Hope Chapel, the smallest of the three movements with approximately 50 affiliated congregations, is a revitalization movement within the International Church of the Foursquare Gospel.

Among Miller's most intriguing insights is his emphasis on the importance of emotional expression and comforting touch in new paradigm churches' worship. According to Miller, participants commonly report feeling "an incredible sincerity" and spontaneous emotional release in the context of communal worship (p. 88). Use of William James to emphasize the foundational importance of religious experience is both a strength and weakness for interpreting this aspect of these groups' success. It raises interesting questions for rational choice theorists who pay little attention to the emotional and relational aspects of religious ritual. However, the resort to James prevents Miller from raising important questions about the affinity of these religious groups for their adherents. Could it be that those who perform emotional labor in the course of doing their service sector jobs find in new paradigm congregations a congenial place where emotions are expressed for their own sake rather than in service of company profit?

Reinventing American Protestantism is written for a general audience. As a result, sociologists will likely be disappointed with Miller's often

suggestive rather than in-depth treatment of theoretical resources. For example, he misses an opportunity to contribute explicitly to the debate regarding supply-side versus demand-side emphases among rational choice and new paradigm theorists. Additionally, his discussion of the extent to which new paradigm churches participate in the "circulation of the saints," documented by Reginald Bibby, raises as many questions as it answers. Finally, Miller differentiates new paradigm churches from fundamentalist, evangelical, or pentecostal/charismatic groups, arguing "members of the Calvary, Hope, and Vineyard movements constitute a 'new paradigm' precisely because they do not fit the traditional categories" (p. 121). As a result of this definitional narrowness and failure to expound on family resemblances between new paradigm churches and other contemporary movements, Miller fails to engage relevant scholarship about charismatic/pentecostal or evangelical movements—such as Mark Shibley's *Resurgent Evangelicalism in the United States* (University of South Carolina, 1996)—that could have thrown additional light on the historical antecedents and contemporary counterparts to these new paradigm movements. Despite its shortcomings, *Reinventing American Protestantism* provides a readable introduction to several rapidly growing churches and will undoubtedly spur spirited conversation about what lessons mainline Protestant denominations should take from the success of these new evangelical groups.

God's Daughters: Evangelical Women and the Power of Submission. By R. Marie Griffith. Berkeley and Los Angeles: University of California Press, 1997. Pp. xi+275. \$24.95.

Susan D. Rose
Dickinson College

God's Daughters is an excellent ethnography that illuminates the stories of women who are members of Women's Aglow Fellowship International (hereafter Aglow), the largest women's evangelical organization in the world, initially inspired and organized by wives of members of the Full Gospel Business Men's Fellowship International. Through an examination of published stories and oral testimonies of charismatic women in Aglow, Griffith reveals the complexity of compromise and the joys, struggles, and contradictions that North American evangelical women negotiate as they make their way in both Christian and secular worlds in the late 20th century. In describing Aglow's organizational structure, expectations, and codes of conduct, and the ways in which women variously accommodate, resist, subvert, and reinforce them, Griffith offers a vivid portrait of the organization and the multidimensionality of Aglow women. In her discussion of the power and role of prayer, Griffith argues that evangelical women's groups parallel other kinds of more feminist consciousness-raising groups where women come together to share sto-

ries, support and comfort one another, affirm sisterhood, and create solidarity.

Aglow members, as both evangelical *and* modern women, simultaneously embrace and resist, criticize, and consume contemporary culture. While they condemn greed and consumerism, Aglow leaders model expensive clothing, a sign both of God's blessing and exemplary behavior and taste, even as they pass the collection plate, exhorting members to sacrifice more to God. Themes of prosperity theology permeate Aglow culture: women are told that their contributions are good investments, while well-coiffed leaders are considered to be "good ads for God" both in the United States and abroad. And international trips and meetings enable Aglow women to travel and experience the world, as I witnessed when I boarded a plane in Seoul, Korea, that was filled with Aglow women who had just visited David (formerly Paul) Yonggi Cho's Full Gospel Church (the largest in the world) and Prayer Mountain.

But this book is about more than this. Griffith examines the themes that emerge in Aglow women's stories: marital discord, childhood abuse, suicidal thoughts, divorce, disordered eating, and the ways in which women's stories integrate the discourse of the recovery movement. In examining Aglow publications from their beginning in 1969 to the present, Griffith notes the shifts in theological thinking and the psychological framing of issues of abuse, blame, remembering, and forgiveness. Although forgiveness remains central, women and children are no longer expected simply to tolerate abuse and to carry the blame and shame of it as theirs. Instead, Aglow prayer meetings and testimonials offer women the sacred space within which to talk of their abuse and their abusers, while blame is attributed to satanic forces and spiritual warfare is invoked to remedy domestic discord and restore domestic harmony. Female submission and surrender to male authority is still strongly emphasized even though there is a perceptible shift among some Aglow leaders to call for "mutual submission" of husbands and wives; it will be interesting to follow the conversation here among Aglow women, the Promise Keepers, and its female counterparts, the Praise Keepers and Promise Reapers.

In the introductory chapter, Griffith presents a solid balance between her own reflectivity about her role as researcher and the stories the women themselves tell. Rather than judging these women or critiquing their choices, she aims to reveal the common ground between evangelical and feminist women and to challenge feminists to take their evangelical sisters more seriously if they want to genuinely recognize and celebrate women's diversity. Griffith takes great pains to emphasize the "common cultural roots of . . . divergent strategies," wanting to "elucidate what is shared among players in the so-called gender wars" (p. 203). While she is clear about her desire to do so, she tends to repetitively assert rather than document that a common ground exists. While she acknowledges and describes contradictions, she does not examine why it is so difficult for evangelical and feminist women to bridge the gender gap that reflects

the varying, and often contradictory, notions of what it means to be a woman.

Evangelical women, like many other American women, struggle with eating disorders, abusive relationships, and self-esteem. But while Griffith details the struggles revealed in the narratives of these women and the similarities among women's struggles (evangelical and nonevangelical alike), she does not address the ways in which various evangelical, feminist, and recovery movements differ in their analysis of and response to cultural constructions of femininity, beauty, or abuse. While Griffith acknowledges that Aglow women, like many of "us," "participate more than they may know in recreating and supporting the very cultural norms they criticize" (p. 16), she emphasizes the similarities of experience at the expense of appreciating the differences in interpretation, analysis, and action.

As a case in point, she refers to Marabel Morgan, evangelical author of *The Total Woman*, as "helping women [to] have better, more 'adventurous' and fulfilling lives" (p. 207) but does not examine Morgan's messages and methods in encouraging women to accept and admire their husbands as they are and to take responsibility for sustaining and enlivening their marriages by becoming "yes women" who are ready to have sex at any time of day. While many female and male evangelical groups are placing greater emphasis on men taking more responsibility in and for their families, they remain self-avowedly patriarchal. Evangelical women, both in the United States and Latin America, may find themselves "transforming traditional sex role hierarchies" (see p. 14), but gender equality is not one of their goals or aspirations. Although Griffith raises important questions, such as: does prayer enable these women "to reenter their daily lives renewed and healed, or does it merely soothe temporarily their anxieties, while binding them more tightly into structures of authority and discipline that give rise to their suffering in the first place?" (p. 79), she does not follow through in answering them.

Griffith calls for a more complex and diverse understanding of women's lives in the late 20th century and one that includes evangelical women. I agree that there needs to be less knee-jerk antipathy and a more complex understanding of the lived experiences of all American women, but I want more discussion of why this is so difficult to achieve. The implicit assumption is that gender transcends class, race, sexual orientation, political, and religious beliefs—an assumption that I am not ready to make.

Evangelical and feminist differences—and those of emerging evangelical feminists that Griffith does not address—are not just a matter of taste or lifestyle: they represent serious and varying understandings of the world, of gender, of equality, and of "choice." Faith is personal and individual, but it is more than that. I want to know more about what kinds of compromises evangelical women (and men) make as they negotiate work, family, the "second shift," and sexuality, and what the respective costs and benefits are. I believe that faith does empower and that Aglow

women feel empowered by collective and individual prayer; the question is, What do they feel empowered to do?

Gracefully written, *God's Daughters* is rich in description and analysis of the stories women tell, the power of prayer, and female solidarity within their circle of faith. But how does the world of Aglow interact with and affect other worlds, other domains? How may Aglow women's choices and agendas affect other women's and girl's choices in relation to sexuality and reproductive health—be that sexuality education, “abstinence-only” education, availability of contraception, abortion, studies of adolescent sexuality, homosexuality? What is not explored is how evangelical women may be empowered to attack (or dismiss) those whom they may think lie beyond their borders: lesbians and gays, pro-choice feminists, and women and children who are impoverished as a result of gross economic inequalities and “welfare reform.” These are significant questions that we all need to examine as we explore how Aglow and other women contribute to the cacophony of voices that are competing to define American reality and morality in the late 20th century.

Splintered Sisterhood: Gender and Class in the Campaign against Woman Suffrage. By Susan E. Marshall. Madison: University of Wisconsin Press, 1997. Pp. xii+347. \$55.00 (cloth); \$21.95 (paper).

Kathleen M. Blee
University of Pittsburgh

When substantial numbers of women mobilized in opposition to the Equal Rights Amendment (ERA) in the 1970s and early 1980s, many observers commented on the paradoxical nature of such politics. Why would women actively campaign against the extension of their rights? How could leaders and activists in such movements justify their participation in public political actions on behalf of restricting women to traditional, private domestic roles? To what extent were antifeminist women activists simply the political pawns of conservative men?

Susan Marshall's book helps us understand the puzzle of women working against women's rights by examining the influential campaign against women's suffrage in the late 19th and early 20th centuries. Drawing primarily on records of the Massachusetts Association Opposed to the Further Extension of Suffrage to Women and a sample of antisuffragist writings, Marshall argues that the campaign against women's suffrage drew its leadership and political strength from conservative elite women who saw the extension to women of the right to vote as a threat to their gendered class interests. In the Massachusetts case, these were women of solid upper-class standing—the Brahmins of Boston society—whose own personal political access was guaranteed by dense networks of well-connected male family members rather than by their own careers or educations. For them, political influence through kinship networks was as-

sured, even if the route to political access was indirect and dependent upon husbands, fathers, and brothers. In contrast, they saw gaining direct access to political power through women's suffrage as threatening their established roles as cultural arbiters and society volunteers and diluting their (albeit indirect) access to the political power held by the nation's richest, most socially prominent men. The inclusion of women in public life, whether through women's suffrage or through progressive reforms that made a place for expert professionals in the provision of social services, threatened the doctrine of separate spheres—the assignment of white women to the private and white men to the public spheres of life—upon which the power base of this privileged, urban elite group of women rested.

Marshall's analysis challenges conventional understandings of women in the anti-women's suffrage movement. Most important, Marshall positions antisuffrage women analytically as discerning, well-connected political actors, not as isolated, manipulated, or uninformed pawns. In so doing, she restores agency to conservative women's politics, underscoring the importance of exploring the roots of such political movements by focusing on the ideologies and social positions of their actors rather than dismissing their politics as auxiliary or as incidental to men's political actions. She also avoids making the issue of gender the sole pivotal issue dividing suffragists and antisuffragists since, as a number of historical accounts have shown, the women's suffrage movement itself drew on conventional as well as progressive views of women's roles. In addition, Marshall's work advances our understanding of those social movements spawned by opposition to social change—often labeled "countermovements"—by pointing to the complicated sets of strategies through which antisuffragists and suffragists contended for power. Those opposed to women's suffrage did not simply react in opposition to the political tactics of the suffrage campaigns; rather, they developed their own proactive strategies in response to perceived opportunities. This is most clear in the analysis of the rhetoric of the antisuffrage movement, which shows a progression over time in how women's roles are presented, from compliant housewife, to scientific homemaker, shaper of the workforce, and independent political actor within a restricted sphere. It is also evident in a fascinating discussion of how the women's club movement served as an organizational base from which women were mobilized into both the suffrage and the antisuffrage movements.

Several themes in Marshall's work are more suggestive than definitive, indicating the need for further studies of the campaign against woman suffrage. Marshall focuses on several local campaigns to try to determine the social characteristics of those who joined antisuffrage organizations or who voted against women's suffrage, but the evidence belies simple explanation. For example, membership in the Maine Association Opposed to Suffrage for Women included not only the expected group of homemakers married to professional husbands but also upwardly mobile single women. Marshall contends that this latter group was searching for

symbolic prestige by allying themselves with a more established elite group, but additional local membership studies are needed to substantiate this conclusion. Likewise, Marshall's argument that women in the anti-suffrage movement acted relatively autonomously from the political direction of their male kin requires additional data to be fully persuasive. Despite these limitations, this volume will be useful to scholars and students of social movements, politics, and gender.

Changing the Wor(l)d: Discourse, Politics, and the Feminist Movement.
By Stacey Young. New York: Routledge, 1997. Pp. ix+249. \$65.00 (cloth); \$17.95 (paper).

Katie King
University of Maryland

Acts of translation from one disciplinary site to another, indeed from one set of interdisciplinarity to another set, are crucial to feminist scholarship. Inherently involved in contests over and about disciplinary values, this scholarship is likely to be on the one hand insightful and enthusiastic and on the other labor intensive and frustrating. Obviously reworked over the decade of the nineties, Stacey Young's *Changing the Wor(l)d* is gracefully elaborated, highly explanatory, and intended as an intervention into assumptions about method, subject matter, and value in the social sciences broadly, and in the fields of government and politics and social movements theory more specifically. It would make a fine text for advanced undergraduates or beginning graduate students, as well as a useful resource for the shelves of feminist theorists and feminist social scientists who want to explain clearly to perplexed colleagues why they depend upon postmodern feminist critiques for insight into their own scholarly projects or the assumptions of their disciplines.

Young works to place "discursive feminist activity" at the heart of inquiry into the U.S. women's movements of the seventies, eighties, and nineties. She uses the occasions of feminist small press publication of particular writings, which she calls "autotheoretical texts," as a wedge into social movements method and theory. She begins with an economic and political description of feminist press activity, drawn largely from interviews with founders and operators of four presses (Kitchen Table: Women of Color Press, Firebrand Books, Aunt Lute Books, and South End Press) and then talks about how the networking activities of the presses reveal their differing political and economic interests. Along the way, she makes very clear how much work feminist editors and publishers put into new authors, into new kinds of writing, and into creative visions of political life. The point of this kind of focus on feminist publishing is to locate it as political activity in both micropolitics between and within presses and macropolitics as economic agents and as ideological agents.

Young then turns to the autotheoretical texts in order to challenge assumptions about identity politics, autobiography, theory, and power. She looks at two pivotal texts, Minnie Bruce Pratt's *Rebellion* and Cherrie Moraga's *Loving in the War Years*, against the backdrop of the anthologies that created one strand of identity politics in the United States in the early eighties. She values the autotheoretical texts as practices in a theory of power that thickly describes "domination and resistance at the level of identity and everyday life" (p. 97). The point of this focus is to redefine what counts as political action, explicitly moving beyond a liberal focus on social policy to acts of "resistance to hegemonic forms of domination" (p. 98).

The second half of the book critiques social science studies of the women's movement and social movement theory and ends with an argument for the use by social scientists of postmodern theories of power, subjectivity, and resistance. The first half of the book is the most obviously original, but the second half is even more thoroughly engaged in the *acts* of interdisciplinary translation that are also a form of "discursive feminist activity." Chapter 4, with its belabored comparison and contrast between women's movement studies by Myra Marx Fereee and Beth B. Hess, Barbara Ryan, Jo Freeman, Ethel Klein and Joyce Gelb, can be read as an exhaustive critique of unintended reinscriptions of liberal ideology but also can be read, along with the subsequent chapters, as a kind of experimental social method that reads feminist social science scholarship as productive misunderstanding and that shifts what can count as "data" in studies of social movements.

Young's book handles feminist controversy in a very understated way, allowing explanation to substitute for debate. But it is precisely because her translations are between fields of power and open to manipulations and misinterpretations that she works so carefully to background the conflicts in which she has been only too embroiled. To the extent that I think there is something one might call interdisciplinary feminist methodology, it is about methods coming into being that center upon acts of translation within and between fields of power. Often one knows such methods when one sees them only in the midst of misunderstandings and struggles, when previously held assumptions are ruptured by micro and macro movements of power. To pay attention to such methodology coming into focus requires a high tolerance for conflict and for beginning again, tasks with emotional, intellectual, and political costs. Young's book is the result of such political work and explores such methods and translations.

Beyond Hierarchy: Gender, Sexuality and the Social Economy. By Sarah Oerton. London: Taylor and Francis. Pp. vii+216. \$24.95 (paper).

Joyce Rothschild
Virginia Polytechnic Institute and State University

In her book, *Beyond Hierarchy*, Sarah Oerton endeavors to apply gender analysis to the study of "flatter" organizations, her preferred term for more participatory organizations. Noting that the second wave of feminism promoted collectivist and cooperative forms of organization as the way to reduce or eliminate hierarchical power relations, her chief research question is whether these recently formed collectivist organizations have succeeded in bringing gender equality.

To address this issue, Oerton chooses a matched sample of some 45 flatter organizations in the United Kingdom, some of single-sex composition and some with mixed-sex compositions. Some of these organizations are in retail, some manufacture, while others provide services. What the chosen organizations have in common is that they are all to some visible degree participatory, cooperative or "flatter" in nature. Oerton interviews only one individual from each organization. This design of the research raises difficulties, to which I will turn later, in the interpretation of its findings.

The first five chapters of the book are devoted to laying out the issues and the relevant literature. Here Oerton correctly notes that most (though not all) of the literature on cooperative and collectivist organizations views their significance to lie in the extent to which they manage to eliminate hierarchies in the workplace and to lessen capital's exploitation of labor by extending to workers both ownership and control of the workplace. Since gender relations are not key in the previous research, Oerton fills an important gap in addressing her work to this issue. Oerton's review of occupational segregation by gender in the general economy is well done, as is her discussion of how gender and sexuality may affect the culture of workplaces.

In chapter 7 Oerton turns to a presentation of her empirical findings. Here the author reports, for example, that even in the flatter organizations that are the subject of her study, women make less money than do men, and they engage in more unpaid extra work than the men. However, because the author has asked only one person per organization what their pay is and has not asked them for the pay distribution of their whole cooperative, her finding concerning pay is difficult to evaluate. In chapter 8 the author argues that the women in the mixed-sex cooperatives and in the women-only cooperatives are marginalized, patronized, and in general treated as more deviant than are the men in these flatter organizations. However, her conclusion that gender and sexuality are the chief basis of perceived deviancy is unconvincing in that it does not take into account that fact that all cooperatives (and by imputation cooperative members) are "deviant" or oppositional in a capitalist economy; indeed,

they often set themselves up to stand as alternative models of how work could or should be done. Predictably, the book concludes that these cooperative and flatter organizations are quite limited in the extent to which they can overcome the inequities in gender relations that are in evidence in more conventional organizations. However, because Oerton has not observed these cooperative groups in action, she cannot speak to whether in fact there are any gender differences in the division of labor, ownership stakes, or control in decision-making processes. Additionally, because she does not compare these flatter groups to the workings of more bureaucratic organizations, either in theory or in practice, she cannot speak to whether or in what ways these groups might alter gender relations in the workplace. The reader is left wishing for more empirical substance on the major question raised by the book.

Based on her respondents' reports, Oerton concludes that some of these groups are so sexualized by the outside world that the women in them are imputed to be lesbian and thus the organizations as a whole are seen as "beyond the pale" or "lesbianized" (her terms). Given that this is one of Oerton's primary findings, it is an oversight that she interviews no one in the environment of these organizations to determine its validity, not in its customer base, nor among its suppliers. Instead, she bases her finding on the reports and perceptions of the few people in these organizations whom she has interviewed, and since this was only one person per organization, the reader is left to wonder whether the perceptions that are at the heart of her conclusions are actually valid. We are left with a disquieting picture: We do learn that the male respondents in her study appear to be more satisfied with what they experience as a less competitive and more humane workplace, and that many of the female respondents feel disrespected in their work and fear that they are marginalized, if not lesbianized, by the outside world. However, since this study contains no observations of the actual practices of these groups, we do not learn whether the men and women in these groups have in fact equal voice in decisional processes. In addition, since outsiders are not interviewed, we do not learn how the outside environment actually views these groups or their members.

In sum, the first half of this book makes a good case for applying gender analysis to flatter or more cooperative organizations, but the second part of the book, the empirical presentation, does not live up to its promise. Because the perspective chosen is subjectivist rather than structural, and because the method chosen limits the study to only the personal reflections of a handful of members on their experience in the co-ops and contains no observations of actual practices or interactions in the co-ops and no observations of people on the outside, the book fails to take account of the objective position of these organizations in the wider economy and fails to appreciate how these organizations may be different than the hierarchical organizations that their own members had hoped to avoid.

Crisis and Political Beliefs: The Case of the Colt Firearms Strike. By Marc Lendler. New Haven, Conn.: Yale University Press, 1997. Pp. x+186. \$30.00.

Judith Stepan-Norris
University of California Irvine

Mark Lendler has written a provocative book. In it he argues that a group of workers became radicalized as a result of a crisis situation that temporarily disrupted normal social relations. He contends that such "moments of madness" enable workers to visualize what previously seemed impossible: challenging their subordination. These moments are "rare and evanescent" (p. 9) and range from revolutions to dramatic strikes. This particular case involves an extended strike at the Colt Firearms plant in Hartford, Connecticut, conducted by United Automobile Workers (UAW) Local 376. The walkout on January 29, 1986, followed a nine-month in-plant struggle and developed into a "dramatic and bitter four-year conflict" (p. 9) that ended in a settlement largely favorable to the union.

According to Lendler, not many strikes have the potential to alter workers' consciousness. This one did because it involved prolonged confrontation, unbearable tensions both inside the plant (*vis-à-vis* management) and outside (*vis-à-vis* strikebreakers), extended hardship, and active support of the union. Lendler's strategy is to follow Barrington Moore's advice by assessing how the workers viewed their lives, how they saw fortune and misfortune, justice and injustice, and how they explained these matters to themselves. His information is drawn from open-ended interviews with strikers, a survey of strikers, newspaper accounts, National Labor Relations Board (NLRB) transcripts, and ethnographic observations.

Lendler's research design attempts to evaluate how Colt strikers perceived American institutions after, and in light of, their strike experiences. Yet he begins his study two years into the strike. His before-after design is thus severely hampered by a lack of an assessment prior to the incident. Also problematic is Lendler's (admittedly) hastily constructed survey of 253 strikers to measure the impact of the incident. In order to show that the Colt workers were radicalized by the strike, he compares their responses (during the strike) to those of respondents from comparison groups. These include samples of other Connecticut union members (p. 78), the national population (p. 89), and workers at one unionized and one nonunionized plant (p. 101). In each of these cases, Lendler assumes (sometimes after minor adjustments) that the samples are comparable, that is, that the Colt workers would have responded in a similar manner in the prestrike days. He finds a greater tendency among the strikers to challenge the idea of the American dream and attributes these differences to their strike experiences.

Although some of the differences he found may be attributable to the

strike experience, a good portion are likely due to his comparisons of unlike populations. Striking Colt UAW members cannot be equated with the general population or even to the blue-collar segments of the sample because both include a range of respondents whom we would expect to be more conservative than the strikers. First, the Colt sample consisted of all unionized workers, while the blue-collar component of the national sample consisted mainly of nonunionized workers (less than 20% of the labor force was organized at the time). Second, the Colt sample is biased toward more liberal prounion activists because it excludes union members who abandoned the strike effort (including 225 who crossed the picket line and 40 who drifted away during the strike). Because the populations are different, it is difficult to determine whether the Colt workers' political responses were in fact more critical of the establishment than those of *like* workers.

Even if we could determine that Colt strikers had more critical views, without a before-strike test it is unclear whether this difference is attributable to the strike experience or to some preexisting condition. Why might we expect the latter? Lendler reports that the Communist-dominated United Electrical, Radio and Machine Workers Union (UE) represented the Colt workers during the 1940s, but that there is no indication that this union left a radical legacy. He arrives at this conclusion based on the fact that he did not hear workers discuss the UE leadership during his investigation. But legacies are about setting patterns of interaction and not necessarily about open discussion of how such legacies were set. If any differences exist, they may have resulted in part from the more militant and progressive tone set by the UE back in the 1940s.

In principle, I am predisposed to agree that prolonged periods of conflict tend to make workers rethink their political views. Lendler's in-depth interviews in fact provide a telling account of how strikers negotiated the events that transpired. If he had placed more emphasis on this rich qualitative data, the book would have been more convincing.

Don't Burn It Here: Grassroots Challenges to Trash Incinerators. By Edward J. Walsh, Rex Warland, and D. Clayton Smith. University Park: Pennsylvania State University Press, 1997. Pp. xiii+292. \$50.00 (cloth); \$17.95 (paper).

Thomas E. Shriver
Oklahoma State University

In *Don't Burn It Here*, Walsh, Warland, and Smith offer an insightful analysis of the anti-incinerator movement. The technology for waste-to-energy incinerators emerged in the 1970s to deal with the growing problem of municipal garbage. With the development of these new incinerators that convert garbage into electricity, trash was potentially transformed from a nuisance to a valuable commodity. Although early

proposals were met with little resistance, by the mid-1980s organized protests were fairly common. By the 1990s a national resistance movement made it difficult for companies to site these incinerators. Walsh et al. examine eight proposed incinerator sites in the Northeast in the early 1990s using survey techniques, interviews, and document analysis. Studying both proponents and opponents, the authors discuss three successful sites and five that were defeated. Results from survey data were inconclusive, but findings from the fieldwork were informative.

The authors identify several key factors distinguishing the defeated from the successful sitings. In communities where citizen resistance defeated proposed sitings, activists had a political rather than legal focus, and they spent much of their time involved in activities such as lobbying, circulating petitions, and endorsing referendums. They developed important support bases with outsiders and were able to exploit the political cleavages among community elites. In contrast, in communities where sitings were successful, activists focused on litigation issues and had minimal outside support. These communities were also more likely to have a political champion campaigning for the incinerator.

Walsh et al. offer several key insights into grassroots environmental resistance. First, the authors illustrate the important link between environmental sociology and social movements analysis. As they point out, "The evidence reveals that each of these protests is really part of a loosely linked national movement that, especially via its grassroots components, illustrates the cross-fertilization of environmental sociology and social movements currently in progress" (p. 43).

Second, the authors discuss an emerging synergy between national and grassroots organizations. Environmental sociologists have emphasized the contradictory goals and strategies of grassroots and national environmental organizations, but the anti-incinerator movement demonstrates that national and grassroots organizations can work together for common goals. Thus, the authors suggest a new synthesis for the future and question the popular sentiment that the entire national wing of the environmental movement is elitist.

Third, Walsh et al. challenge traditional social movement notions of networking. Social movement analysts (in particular those operating out of the resource mobilization tradition) have emphasized the importance of preexisting networks in social movement success. This research indicates that preexisting networks may be overemphasized, as they show how newly formed networks based on suddenly imposed grievances can be equally effective.

Fourth, the authors contribute to our understanding of framing in social movement analysis. While contemporary social movement analysts have emphasized the extent to which master frames might be adjusted to include more supporters, Walsh et al. show how activists expanded their master frames to reach others sympathetic to their struggles. In addition to adding supporters, this strategy effectively allowed activists to

"neutralize the NIMBY label by expanding their base of support beyond the backyard area" (p. 259).

While the book makes a number of important contributions to social movement literature, I take issue with the distinction between equity and technology movements. This dichotomy implies that technology movements are not based on equity/justice. However, one of the organizing themes of the anti-incinerator movement is fairness of siting decisions, clearly an example of environmental justice and equity. The authors argue that these new technology movements are based on suddenly imposed, rather than long-standing, grievances. Again, I disagree. Those social groups targeted by newer technologies are the same ones (e.g., minority, lower class) who have suffered from other forms of injustices. In addition, the master frames used by these newer "technology" movements have been provided, in part, by older "equity" movements. In communities facing environmental pollution from newer technologies residents are just as concerned with justice and equity as the Civil Rights movement was 30 years ago. I am not convinced that technology movements possess characteristics that set them apart from older equity movements.

This work represents one of the first systematic attempts to analyze the anti-incinerator movement in the United States. The authors have a rich and unique data set. The book is well written and makes important contributions to environmental sociology and social movement analysis.

The End of the Peasantry: The Rural Labor Movement in Northeast Brazil, 1961–1988. By Anthony W. Pereira. Pittsburgh: University of Pittsburgh Press, 1997. Pp. xxi+232. \$45.00 (cloth); \$19.95 (paper).

Nancy D. Lapp
University of California, Los Angeles

As a case study on the northeastern state of Pernambuco (not the northeast as a whole, as the title implies), this book gives a richly detailed account of the rural labor movement in that state. Using surveys and interviews primarily conducted in 1988, as well as Brazilian archival and secondary sources, Anthony W. Pereira describes the rural unions of the sugar zone of Pernambuco. Firsthand observations of strike activity and colorful anecdotes enliven the account. Where Pereira does make explicit comparisons with other regions, the results prove particularly interesting.

The case study depicts the emergence of a reinvigorated rural labor movement shaped by the dual influences of the Brazilian military regime and the capitalist transformation of the countryside—a transformation that led to the "end of the peasantry," (i.e., the marginalization of subsistence cultivators). The resultant union structure yields an answer to one of Pereira's key questions: Why, despite an avowed commitment to land reform by the national labor federation, did local labor unions dedicate

their efforts instead to wage increases and the enforcement of union contracts? The answer is that these were the issues that concerned wage workers and small farmers, the bulk of the union members. Rural labor leaders dependent upon election by a union membership devoid of peasants had little incentive to pursue land distribution. Pereira also explains that when the military government created a welfare system administered by local unions, local union leaders subsequently dedicated more time to providing health, dental, and retirement benefits than attending to other union business such as monitoring employer compliance with the union contract.

Brazilian land reform proponents anticipated significant advancements in land redistribution in the 1980s. This book helps explain why a large-scale land reform did not materialize and why the most dynamic agitation for land redistribution has been conducted by groups such as the Landless Workers' Movement (MST) rather than through the auspices of official unions. Pereira also provides a succinct and welcome summary of recent land reform history in Brazil, although the book is not primarily about land reform. In the book's strength, however, also lies its weakness. The case of Pernambuco cannot be generalized automatically for all of Brazil; as Pereira himself acknowledges, the state's rural unions are not "typical" (p. xvii). Here the emphasis on Pernambuco is felt the most, since the MST has been more active in other regions of Brazil. Furthermore, Pereira tantalizingly notes that neighboring northeastern states experienced more land reform results than Pernambuco, without accounting for this variation in outcomes.

While focusing on rural labor organization, Pereira probes the relationship between societal forces and state control, particularly during democratic transitions. Herein lies his main theoretical contribution. The book complements work by Frances Hagopian on Minas Gerais (*Traditional Politics and Regime Change in Brazil* [Cambridge University Press, 1996]). Both detail how methods used by the Brazilian military regime to shore up political support outlived the regime itself and resulted in important (and mostly negative) consequences for the new democratic government. More than others, Pereira makes explicit the complicated interactions between groups in society and the Brazilian state. In this relationship state institutions shaped union behavior, but the unions also carved out their own victories in their battles with employers and the state.

The End of the Peasantry touches on themes of particular interest to the student of Brazil but should appeal to anyone interested in labor movements and democratization in Latin America. Insightful and succinct, the book provides a wealth of new information and is a worthy theoretical contribution.

Peasant Protest and Social Change in Colonial Korea. By Gi-Wook Shin. Seattle: University of Washington Press, 1996. Pp. xiii+234. \$40.00.

Hagen Koo
University of Hawaii

Peasant Protest and Social Change in Colonial Korea is one of the finest books written on some crucial features of Korea's modern transformation in the 20th century. Shin skillfully weaves together theory and history to present an extremely interesting and insightful account of peasant struggles in contemporary Korea. Shin's theoretical sophistication, his methodological acumen, and his lucid style of writing make this book an exemplary historical sociology study. It makes a great contribution not only to the literature on Korea and East Asia but also to comparative historical sociology.

The study investigates peasant struggles in Korea from the 1920s through the late 1940s and seeks to demonstrate how important a role Korean peasants played during this turbulent period, not simply as victims of colonial exploitation, but as active agents of historical transformation during and after colonialism. Following a respectable tradition in historical sociology established by such scholars as Barrington Moore, Jr., and Robert Brenner, Shin looks at agrarian conflict as a major causal factor in explaining diverse paths to modernity. The main thesis of the book is that peasant struggles during the Japanese colonial period played a critical role in weakening the power of the Korean landlord class and thereby in providing the structural condition for postwar land reforms in the late 1940s. Looked at from the other end, postwar peasant radicalism in Korea had its historical root in diverse forms of peasant struggles during the colonial period and in the changes these struggles had brought about in the rural class structure and in peasants' political consciousness.

Shin first introduces a thoughtfully worked out theoretical framework, which emphasizes the multifaceted and complex nature of peasant protest. Following this, he proceeds to analyze diverse forms of peasant struggles characterizing three periods of colonial rule: the tenant protests of the 1920s, the "red peasant union" protest movements of the 1930s, and the "everyday forms of resistance" in the 1940s. Shin rejects a simplistic thesis of colonialism causing pauperization, which in turn causes social protests. Instead, he argues that the effect of colonialism was much more complex than commonly assumed, producing a differentiated class structure rather than a sharp polarization between landlord and peasant classes. Using a wealth of statistical data and secondary materials, he demonstrates that predominant forms as well as dominant actors of peasant protests varied across different periods, and also across different regions, reflecting different constellations of class interest and class power.

Several important conclusions are drawn from a careful historical analysis, a major one of which concerns the effect of peasant protests. "Troubled by increasing rural unrest, the colonial government undertook legal

measures to protect tenant rights and limit landlord power—for example, the Tenant Arbitration Ordinance of 1932 and the Agricultural Lands Ordinance of 1932” (p. 74). Such a shift in colonial policy led to the weakening of landlord class power and inadvertently encouraged more politically oriented peasant protests in the following decade. Shin regards the undermining of landlord power during this period as a critical precondition for the relative ease with which postwar land reforms were carried out in Korea. He also contends that this colonial policy shift (unfavorable to the landlord class) forced many landlords to turn to industry as an alternative investment for their capital, siding with the argument of the “colonial origin of Korean capitalism,” advocated by Carter Eckert, Dennis McNamara, and others.

A major concern of Shin’s analysis is with establishing the linkages between colonial-period peasant protests and postliberation radicalism. Combining secondary data with statistical analysis, he demonstrates the validity of the proposition that postwar radicalism, like the 1946 uprisings that started in Taegu and spread across the countryside, had its historical base in peasant struggles during the colonial period. For instance, the communities that participated actively in the 1946 uprisings were likely to have had extensive experience in tenant protests in the 1930s. His data analysis provides a firm empirical base for his class-based explanation of postwar peasant radicalism in Korea: “First, peasant radicalism in liberated Korea did not materialize overnight or without presage. It was not simply a result of agitation by communists as official accounts alleged, or by returnees from Japan and Manchuria with radical ideas as Cumings argues. Nor did Japanese defeat in the Second World War simply provoke radicalism. Instead, it emerged from peasants’ long development of both consciousness and a sense of efficacy through participation in colonial protest and resistance, particularly tenancy disputes” (p. 171).

This is one of the most nicely organized and clearly written books I have read recently. I see hardly any flaws in the book. Apart from its great contribution to the accurate understanding of the nature of peasant struggles during the colonial era and their implications for the later period, this book has clearly demonstrated the utility of the class perspective in understanding historical transformation. And unlike some dry, structuralist analyses of class transformation, Shin’s class analysis brings seemingly powerless people (or agency) to the center stage of historical drama and shows beautifully how ordinary people shaped history through their continuous struggles for a better life.

Taking on Goliath: The Emergence of a New Left Party and the Struggle for Democracy in Mexico. By Kathleen Bruhn. University Park: Pennsylvania State University Press, 1997. Pp. xiv+365. \$55.00 (cloth); \$22.50 (paper).

Kevin J. Middlebrook
University of California, San Diego

Kathleen Bruhn's study of the Party of the Democratic Revolution (PRD) makes significant contributions to our understanding of democratization in Mexico and the challenges of party building in contemporary Latin America. She provides an exhaustive account of the PRD's origins and early evolution, showing how a major schism within the hegemonic Institutional Revolutionary Party (PRI) in 1986–87 led to unexpectedly broad popular mobilization behind Cuauhtémoc Cárdenas's presidential candidacy in 1988 and the PRD's formation as a new left-of-center party in 1989. By detailing the PRD's struggle to survive in an authoritarian regime capable of deploying immense institutional resources (including significant repression) against political challengers, Bruhn provides valuable insight into a crucial period in Mexico's political development. Her book makes an especially important contribution to research on the Mexican Left.

Bruhn's conceptual framework distinguishes between the conditions surrounding party emergence and the requirements for successful party consolidation. Although emergent political movements or parties can mobilize substantial electoral support by surprising their opponents and by highlighting the charismatic appeal of their founders, party consolidation depends upon "the construction of stable norms and expectations for mutual cooperation, decision making, and conflict resolution in five key relationships: (1) among activists; (2) with allies; (3) with civil society organizations; (4) with other parties; and (5) with the state" (p. 13).

In the first instance, the success of a new party depends upon its ability to detach supporters from established parties, capitalizing on such matters as changes in existing parties' social bases, popular discontent with political incumbents' economic policies, or public concern about corruption. Over the longer term, however, Bruhn argues that a new party's greatest challenge is to "reattach" the loyalties of potential voters and activists. This is a particularly daunting task in an authoritarian context in which the new party, at least initially, can offer few material or policy rewards to supporters. In the case of the PRD, Bruhn shows how some initial advantages, including Cárdenas's charismatic authority and the breadth of the *cardenista* coalition, later became serious obstacles to PRD consolidation. These problems—coupled with national economic recovery in the early 1990s, the Salinas de Gortari administration's (1988–94) use of high-profile antipoverty programs to restore popular support for the regime and undermine PRD strength, and the PRD's own tactical

errors—largely accounted for the PRD's comparatively poor electoral performance in the early 1990s.

Following a prominent recent trend in the field of comparative politics, Bruhn goes to some lengths to identify ways in which different institutional arrangements shaped the emergence and consolidation of the PRD. For example, she demonstrates effectively that electoral reforms that jeopardized the future of several small "parastatal" parties (those historically aligned with the governing PRI and highly dependent upon state financial and political subsidies) and that reduced their value to the PRI as potential allies, provided strong incentives for these parties to back Cárdenas's 1988 presidential candidacy in the expectation that his long political coattails would guarantee them sufficient votes to preserve their legal status. Similarly, Bruhn shows that struggles over the composition of the party's proportional representation electoral lists heightened internal factionalism.

But at other times, Bruhn somewhat exaggerates this line of analysis. One example is her argument that "presidentialism encouraged policy opponents to form a new party by limiting alternatives within the system for influencing policy against the will of the president" (p. 209). The highly centralized brand of presidentialism that characterized Mexico's authoritarian regime in the 1980s and early 1990s may have had this effect. However, it is less certain whether this is an inherent consequence of presidential systems. Nor is it clear how one should weight this institutional factor in comparison to other elements (such as the impact of prolonged economic crisis and neoliberal structural adjustment on Mexico's heterogeneous governing coalition) that contributed significantly to the rise of a new political option on the left.

Bruhn's study is based on numerous interviews with party leaders and extensive personal observation. Nevertheless, her book offers relatively little insight into decision making within the national PRD, a limitation that sometimes hinders Bruhn's capacity to make sense of party strategy. For example, in contrast to her sophisticated treatment of the *cardenista* coalition's strategic options in the immediate aftermath of the fraud-ridden July 1988 presidential election, Bruhn can only hypothesize that the need to preserve credibility with their social movement constituencies compelled senior PRD leaders to take actions (such as meeting with Subcomandante Marcos, public leader of the *zapatista* rebellion in the southern state of Chiapas) that, at least in the view of many detractors, associated the party with social radicalism and violence.

If only because it was a central axis of debate within the PRD well into the 1990s, the tension between preserving the party's social movement base (and Cuauhtémoc Cárdenas's apparent predilection for conducting politics in a social movement style) and building a more conventional party organization would have merited even more attention than Bruhn devotes to it. Similarly, although Bruhn documents the factional tensions that arose from efforts to promote internal party democracy, she might have placed greater emphasis on the strategic value that internal

democracy held for preserving members' loyalty under very difficult conditions.

The use of more detailed case studies drawn from Bruhn's extensive field research in the states of México and Michoacán might have successfully taken the reader "inside the PRD" on such questions. This would have both enlivened Bruhn's discussion and perhaps revealed more fully the debates and tensions that continue to shape the PRD's evolution. Closer attention to the party's shifting internal dynamics, especially at the national level, might also have laid a firmer foundation for understanding the PRD's remarkable resurgence in 1996 and 1997 and its considerable impact on Mexico's accelerating political transition.

Urban Protest in Seventeenth-Century France: The Culture of Retribution. By William Beik. New York: Cambridge University Press, 1997. Pp. xiii+283. \$59.95 (cloth); \$19.95 (paper).

John Markoff
University of Pittsburgh

For the past couple of decades, historians have devoted considerable attention to people who commanded no states or armies, wrote no famous books, and left no splendid houses to their heirs. The world of the "ordinary people" of the past has often been terribly difficult to reconstruct from the inadequate documents in which their acts or words are recorded—often by administrators seeking taxes or conscripts and police or judicial authorities interrogating detainees. There are great gaps in our knowledge, some of which will never be filled.

One extremely important area has been the study of popular engagement in conflict. For Western Europe before the democratizing revolutionary wave of the late 18th century, we now know vastly more than we did a generation ago about the forms of collective action, the circumstances in which such actions occurred, and the sorts of people who participated. And while there is a great deal still to be done on those matters, social historians are beginning to broach the next big question: with what effect? The question requires that we ask not merely how those down below responded to the acts of those on high, but how those on high altered their activities in response to popular action. And it requires that we ask not merely how those below attempted to live their lives within some institutional order, but to ask whether popular action may have played a role in shaping that institutional order.

Such issues, important at all times, are particularly challenging for the centuries before modern democratic institutions and modern social movements provided regular channels for those down below to pressure those above. William Beik's impressive book on popular challenges to elites in the cities of 17th-century France is an important contribution to these discussions. His careful sifting of the evidence on some 20 cities demon-

strates the pervasiveness with which the urban elites of the so-called age of absolutism were keenly aware of the potential for serious disruption presented by the political culture and organizational traditions of urban plebeians. With great frequency, the 17th-century French state was attempting to impose new demands on its townsfolk, particularly taxation, and, in the view of the central government, it was among the responsibilities of the town's governing bodies to keep order in the face of potential resistance to those demands. Locally, however, the forces of order might be little more than a militia from which those of wealth and influence would be managing to absent themselves and whose plebeian participants might well refuse to defend the tax agents against concerted action by organized groups of artisans or the stone-throwing women and children who figure with great frequency in eyewitness accounts and post hoc inquests.

Beik convincingly shows that early in the century, various elite bodies and individuals often found it difficult to engage in concerted action in the face of challenge from below. Royal judges, tax farmers, princes in command of troops, urban militia commanders, and city fathers often entered crises posed by popular mobilization at odds with one another. At the same time, and particularly strongly developed in Beik's sifting of the documents, the urban popular culture readily legitimated acts of retribution against those held to betray the popular interest by supporting the agents of new taxation.

Beik insists that unity was problematic, both for elites and for plebeians. Around the modal pattern of conflict, Beik takes up instances in which some of the elites themselves acted in concert with popular forces. His repeated demonstration of urban elements that supported, as well as those that opposed, the outsiders is an important corrective to romanticized fantasies of a united people. And he tellingly shows how the most effective resource of those down below was their capacity for disruption.

As the century advanced, the elites, under pressure from the centralizing monarchy, acted with more effective coordination over wider geographic areas. But Beik suggests that such increasingly effective elite action was beginning to produce an increasing propensity for geographically widespread resistance of socially more disparate groups, all now threatened in common by that more effective state.

The impressively marshaled evidence suggests many questions. How were the women and children organized who so frequently appear not merely in the ranks of, but in the forefront of, resistance to tax collectors? Why was even the most radical of the French urban collective actions—the effective if temporary seizure of power in Bordeaux in midcentury—so limited in the development of democratic ideas compared to contemporary developments in England? What role, if any, did such traditions of resistance play in the late-18th-century explosion? How are 17th-century French urban popular politics like, and unlike, that of the previous century, and like and unlike that of other places in Europe—a question treated in an important new work by Wayne te Brake (*Shaping History:*

Ordinary People in European Politics, 1500–1700 [University of California Press, 1998]? And, when we consider both the degree to which challenges could disrupt elite policies and the extent to which collective punishment could bring real hardship to rebellious towns for years, what was the balance sheet of gains and losses for urban protestors before the era of modern democracy?

With great economy, William Beik has surveyed what is known of 17th-century French popular urban politics, has added much detail on many towns, has developed some important new interpretations (and important corrections, emendations and challenges to previous scholarship), and, in so doing, suggests the many matters that remain to be explored.

Radical Artisans in England and France, 1830–1870. By Iorwerth Prothero. New York: Cambridge University Press, 1997. Pp. xvi+424. \$74.95.

Richard Biernacki
University of California, San Diego

In the past two dozen years, the field of 19th-century labor and social history has drawn contributions from a score of prominent historical sociologists and theorists. They have used this field as a site to develop new understandings of how language and tradition steer the development of radical movements. In this well-researched book, Iorwerth Prothero quietly rebuts a startling amount of their accumulated wisdom.

Prothero mounts his challenge by comparing plebeian movements for political and labor reform in England and France from 1830 to 1870. His strategy is sensitive to current skepticism about the nation-state as a unit of comparison. Instead of treating "England and France as two distinct blocks on which to construct national generalisations," he uses comparison to multiply the instances of popular movements and to increase the variety of political contexts in which they emerge. The movements Prothero scrutinizes include assemblages leading to armed disorders, as well as durable trade unions, political parties, co-ops, and self-education groups. He asks how the organizational forms, strategic agendas, and popular support of each kind of activity varied by local setting and by the macropolitical climate. England and France offer matching evidence, Prothero explains, because they were the first nations to begin industrialization. They shared the experience of greatly expanding employment of artisans, who provided the mainstay of the organized movements for change and reform.

From E. P. Thompson onward, preeminent historians have illuminated the course of these artisan-backed movements by examining the movements' cultural origins and long-term intellectual progression. Prothero breaks up these developmental chronologies to refocus on brief and dis-

crete political conjunctures. For example, he sees workers' violent resistance in Paris during June 1848 and in British cities such as Burnley during 1842 as immediate responses to the withdrawal of public relief for the poor. Situational variables directed the course of each kind of movement undertaking. In Britain, energy invested in cooperative societies rose promptly after great defeats of Chartist agitation, especially after 1839, 1842, and 1848. Likewise, in France, labor radicals used co-ops as an outlet after 1848 when political setbacks were conjoined with partial freedom to associate. When repression of labor associations became most severe, however, educational, cultural, and convivial activities became more dominant. Since many of his major chapters center on a specific kind of movement undertaking, Prothero can identify cross-national parallels in the relation of each undertaking to the larger field of radical organizing. Once Prothero has "controlled," so to speak, for the immediate political conjuncture, he finds that the differences between England and France in the strategy and form of each type of radical organization are unexpectedly marginal. Nationally distinctive ideologies no longer serve as an explanatory key.

To underscore similarities between the ideas of artisans in England and France, Prothero breaks the movements' worldviews down into constituent themes. At that level, ideologies as diverse as those of the Icarians and Proudonists in France or of the Owenites and O'Brienites in England become members of a family of ideas Prothero terms "radical." The movements all emphasized rationally organized democracy as a cure for both economic and political inequity. They ennobled labor as a foundation of society. Creeds labeled "socialist" enjoyed wide acclaim in France, but an equivalent emphasis on the social ownership of some types of property and on harmonious production appeared in branches of English radicalism.

Prothero's work ought to stir controversy about the goals and tactics of comparison. When does the discovery of national similarities say more about the generality of the investigator's categories than about the motives and causes at work in history? Prothero constructs the political situation in each country almost as if it were an exogenous given that structures social movements rather than as a human creation issuing from those very movements. Does his picture of underlying national similarities face contrary evidence from the Paris Commune, the event that brings his timeline to a close? When are extreme political conjunctures, such as the French revolution of 1848 and the Commune, treated best as local coincidence and when as expressive in part of underlying, nationally distinctive cultural processes? Some readers will ask what we sacrifice in historical understanding when we handle these traumatic, epoch-making upheavals in Paris as social-scientific "cases" equivalent to any other urban disorder in France or Britain. Prothero is not the only researcher to have put these issues of comparative reasoning on the table. But this important book should prove crucial for their discussion.

Return to the Western World: Cultural and Political Perspectives on the Estonian Post-Communist Transition. Edited by Marju Lauristin, Peeter Vihalemm, Karl Erik Rosengren, and Lennart Weibull. Tartu, Estonia: Tartu University Press, 1997. Pp. xvi+387.

Joan Löfgren

Estonia, a small country on the eastern shores of the Baltic sea, has been touted as the success story of postcommunist reforms since regaining independence from the Soviet Union and is now being included in the initial group of applicant countries to start European Union (EU) membership negotiations this year. In this anthology, Estonia's postcommunist transition is explored in terms of Westernization, more specifically value changes, democratization, and the role of the media. A welcome addition to the transitions literature, the book draws on an enormous body of survey data and helps to bridge the gap between sociological theory and postcommunist cases. It is the first major publication of the Balticom Project, a joint Estonian-Swedish research program, reflecting research carried out from 1991 to 1996 and based on longer-term collaboration. One of the editors, Marju Lauristin, was herself a key figure in the Estonian transition, at one time serving as minister of social affairs in the newly independent state.

Only a few of the arguments by the book's contributors will be critiqued here. First, Karl Rosengren and Marju Lauristin provide helpful theoretical perspectives for interpreting change in Estonian society, including classic notions such as agency and socialization. Lauristin also argues, drawing on Samuel Huntington's recent work, that the most decisive role in the collapse of the Soviet Union was played by the civilizational conflict between the Russian-Soviet Empire and the Baltic countries and other East European nations that looked to Western traditions of individual autonomy and civil society. The Soviet system was "orthodox civilization in its new Soviet-Russian form" (p. 29). How important is this distinction for overall arguments on Westernization? Does defining Estonia as "Western" depend on assuming Russia (the Other) is not Western? The degree of Westernization in Estonia is outlined by Lauristin according to certain indicators, for example, attention to personal success and self-expression (p. 30). This still leaves open, however, the question of Westernization in Russia, which, due to increasing regional and generational differences, cannot be assumed as culturally monolithic. Young Russians in St. Petersburg may in fact have more in common with their Estonian counterparts than with their own parents—Where is the civilizational divide to be seen in that case?

Essential historical background is also provided (chaps. 4–5) that emphasizes the early history of the region and the events of the last decade; oddly, the Soviet occupation period is practically absent and would have been helpful for readers unfamiliar with the region. The study compares Estonian worldviews to those in Finland and Sweden using survey data

measuring international orientations in those societies (chap. 6–8). While the Nordic countries as a whole are considered very important for Estonia, the interest is not quite as mutual; this is not surprising, given Estonia's relative position as a small power, even in a region of small powers.

Several contributions focusing on democratization in Estonia (chaps. 9–11) illuminate aspects of the transition overlooked by other studies, such as political culture and the role of the media. Peeter Vihalemm, Lauk, and Lauristin have compiled data on the media spanning almost a decade and conclude that the Estonian media played a dual role in transition: as tool of ideology and as preserving cultural space. Weibull and Holberg compare "two young democracies" (Estonia and Uruguay) with an "old one" (Sweden), concluding that Uruguay, with its long tradition of multiparty politics, and relatively short military dictatorship, is in fact closer to Sweden as a democracy than to Estonia. The comparison is at times confusing and the reader is left wondering why Uruguay was chosen as a case, but some interesting points are raised in the discussion.

In the final chapters (12–14), value changes in Estonia and Sweden are outlined. Lauristin and T. Vihalemm, drawing on Inglehart et al., point to a potential cultural contradiction in the transition, in a sense between the winners and losers: "Economic growth, openness to the West, and rapid generational replacement should promote post-materialist values, whereas economic and social constraints and the growing insecurity of people create a basis for strong materialist orientations" (p. 250). They also analyze value changes in the ethnic Russian population in Estonia, drawing on an impressive set of survey data from the project and from other sources (R. Rose, Baltic Barometer Surveys, Univ. of Strathclyde; A. Kirch, M. Kirch and T. Tuisk, Institute for International and Social Research, Tallinn). They note increasing individualism among both ethnic Russians as well as among Estonians, suggesting that the Westernization process in Estonia is bringing about common value changes in society across ethnic lines. A perhaps more worrying trend would be a growing divergence between the winners and losers of the transition, reflecting divergent values.

In the brief "coda" to the book Rosengren and Weibull summarize some key points, concluding that (1) Estonia somehow maintained elements of Western culture throughout the occupation and passed those sensibilities down through generations, (2) there are clear signs that the value structures of ethnic Estonians and Russians are converging, and (3) in the Baltic region, differences remain between value structures in Estonia and versus other Western or Northern countries. All of these suggest fruitful research questions for further research.

Return to the Western World is essential for Baltic specialists and can be highly recommended to a broad audience in sociology, political science, journalism, and East European/post-Soviet area studies. It is ambitious in being comparative, which helps overcome the exceptionalism that has plagued Soviet area studies, and in presenting an extensive compilation of survey data on Estonia, replete with tables and charts. It is

highly readable and accessible to a generalist audience. An overall shortcoming in the book is its lack of a comprehensive conclusion. The editors seem to have intended this as a collection of works thus far resulting from the project, but uniting them more thoroughly would have made the book even more satisfying.

Money Politics in the New Europe: Britain, France and the Single Financial Market. By Daphne Josselin. New York: St. Martin's Press, 1997. Pp. xviii+235. \$69.95.

David Knoke
University of Minnesota

In 1992, after seven years of planning, negotiation, and political maneuvering, the European Union (EU) had forged a single financial market that essentially removed existing barriers to transnational services in credit, investment, insurance, and capital. This intricate liberalization process and its impacts on the economic and political institutions of Europe's two preeminent financial centers, London and Paris, make up the substantive core of Daphne Josselin's remarkably detailed reconstruction. But her larger analytic agenda is to demonstrate how a policy network perspective can illuminate the complexities of national and EU politics as that rough continent slouches toward political integration.

Josselin weaves her analytic framework from two theoretical threads: (1) the interest group intermediation debate among European political scientists such as William Coleman, David Marsh, and R.A.W. Rhodes over the explanatory efficacy of contrasting pluralist-corporatist and policy community/issue network ideal typologies and (2) the quantitative structural network approach developed by political sociologists, such as Edward Laumann, Franz Urban Pappi, and me. Each perspective emphasizes distinctive substantive questions, key political actors, and formal methodologies, with divergent explanatory capacities and glaring gaps. Josselin favors a composite macrotheoretical framework that places theoretical concepts characterizing domestic policy networks along a continuum ranging between the corporatist political community—a dense, centralized network dominated by the state and a few power interest groups—and the pluralist issue network, consisting, of weakly connected, highly specialized organizations with limited resources and narrow interests. The empirical punch of Josselin's research comes from her excruciatingly documented investigation of where along this scale the French and British cases fell, and the EU-level network arose, during the transformation of the financial services sector.

Her multimethod project blended archival documents, directory listings, press clippings, parliamentary submissions, and interviews with formal network analysis of issue-interest and communication matrices for the 34 British and 18 French organizations that form the core sets of each

national policy domain. The result is a rich stew of thick description, historical narrative, case studies of three major protracted negotiations, and spatial maps revealing the differential responses of the two national policy systems to the 1986 Financial Services Act and their connections to the Byzantine world of the European Commission.

Although the conventional wisdom had depicted both France and Britain as "extremely low with respect to corporatist features" (p. 23), Josselin uncovered divergent networks as the financial services reforms proceeded: "In the French instance, the community appeared small, tight, and largely revolving around the state. On the side of the British, the network was larger, more elusive on the investment segment and altogether more fragmented and evolutive, but still characterized by precise rules of exclusion, and by what seemed a more equal power-sharing" (p. 67). But state-interest group relations in the British network appeared to tighten, albeit unevenly across subsectors, as negotiations proceeded toward implementing the single market terms. Across the Channel, the already highly centralized French state-interest group old-boys' informal network effectively excluded consumer and labor union concerns. These impressions are reinforced by smallest-space communication configurations showing the British system somewhat balkanized among subsectors, with both the Treasury and Bank of England outside the core ring occupied by the Department of Trade and Industry surrounded by numerous claimant groups. In France, the Trésor and National Employers Council sit like spiders in splendid isolation at the very center of the web.

The question, "Are certain features of domestic policy networks preconditions to the exercise of political influence at the supranational level and, if so, which ones?" (p. 10), is not easily answered. Only one of Josselin's three in-depth case studies of network activation involves the promotion of national preferences at the European stage (the EU's position on the General Agreement on Trade in Services). The number and complexity of interorganizational interactions in the emergent EU polity are so intricate as to defy detailed examination. The EU's financial services subpolity seems to lack the cohesion and consensus exhibited by its agriculture, technology, and chemical industry policy communities, possibly because of the "quintessentially competitive and fragmented nature" of the finance sector (p. 180). Josselin's impressionistic conclusion is that, given such a loosely integrated supranational issue network, Britain's multipolar financial services network seems better situated than the rigidly centralized French network in mobilizing support necessary to promote its views. A more definitive assessment of the relative efficacy of national policy networks operating within a transnational political system will require a more ambitious, comprehensive, and well-funded research effort.

Daphne Josselin's protean exploration of money politics in the EU only begins to hint at the potential payoffs from taking a policy network perspective. She argues that the approach is "better suited for description and comparison than for ground-breaking explanation" (p. 193). But, by focusing her attention so extensively on answering substantive questions

and squeezing her observations into preconceived analytic boxes, she missed an opportunity to bind her empirical bricks with theoretical mortar and thus erect a truly innovative structure. Determining whether empirical political systems conform more closely to pluralist, corporatist, community, or issue-network ideal-types is less creative than would be articulating testable theoretical propositions about the conditions under which interorganizational relations emerge, network structures evolve, strategic actions shift, and policy outcomes change. This valuable book hints at, but fails to realize, the full power of policy networks to explain the growing integration of national and global political economies.

Modernization and Postmodernization: Cultural, Economic, and Political Change in 43 Societies. By Ronald Inglehart. Princeton, N.J.: Princeton University Press, 1997. Pp. x+453. \$60.00 (cloth); \$18.95 (paper).

Avner Offer
Nuffield College, University of Oxford

This is the latest installment in a long-running research program on post-materialism that goes back to the early 1970s. The central ideas are simple: as people grow wealthier, they experience diminishing returns from material goods, such as food, clothing, and motor cars, and develop an appetite for nonmaterial goods like the environment, personal relations, and civic values. This "postmaterialist" orientation increases among younger cohorts, because they have grown up in a more affluent environment, and they have not experienced the traumas of economic depression and world war. Formative influences are important: postmaterialism is nurtured by a sense of security in childhood and adolescence. The original prediction was confirmed in several European countries with a four-item questionnaire. It has attracted a great deal of criticism, mainly on two lines: the definition of "post-materialism" was too vague, and the theory (which appeared plausible in the context of the 1960s) did not square with the observed political and social trends.

Ingelhart has persisted, and this book (following two earlier ones in the 1990s) extends the analysis to a couple of large international surveys of values from the early 1980s and the early 1990s, covering some 43 societies in all. As before, a small number of questions (12 this time) are used to uncover a materialist-postmaterialist spectrum. This is transformed into a single point on a two-dimensional field, with a horizontal security → quality of life axis, which denotes materialist → postmaterialist attitudes, and a vertical traditional → secular authority axis. Among the 43 societies, several clusters emerge. There is an overall upward diagonal thrust, from Poor (India, South Africa, Nigeria) at the southwest end to Nordic (Scandinavia and Holland) at the northeast corner, a trend which confirms the progression from materialism to postmaterialism on the ladder of affluence. There is an ex-communist country cluster, low

on security and high on secular authority; the Catholic countries of Europe (Italy, France, Austria, Spain) are midway on the diagonal, somewhat lower than their affluence would imply. The English-speaking world is higher on quality of life and lower on secular orientation. A good deal of more detailed analysis of particular issues and attitudes elaborates these findings. Progress is not immutable: economic recessions can deliver setbacks, and some countries resist the trends. The time elapsed since the first surveys confirms that younger cohorts are more postmaterialist and remain so as they grow older. The displacement of older cohorts by younger ones drives up the proportion of postmaterialists, making them a majority in some countries. The radical student cohorts of the 1960s are now middle aged, in positions of influence and authority, and more able to affect public discourse and priorities.

The theory accounts for several social and political trends, for example, the decline of authority and the realignment of politics away from a left-right orientation toward issues-and-attitudes. GNP measures are biased toward material goods, and the countries with the highest economic growth are also the most materialist. The decline of growth in the most advanced societies confirms their shift into postmaterialism. The enterprise is dazzling in its suggestive richness, and unmatched as a sustained analysis of recent social and cultural development. By sheer determination, energy, persistence and technical skill, Inglehart has challenged history to prove him right. But gratitude for this major illumination is qualified by a few doubts. That is good: Inglehart himself insists that postmaterialism is not the end of the road for cultural and social change and that reality is complicated by a great many other factors.

One problem that has dogged the program from its inception is the correspondence between survey questions and analytical categories. The shift from four to 12 questions is an improvement. Yet the technical procedures of clustering are not sufficiently transparent. Later in the book, when discussing multivariate procedures, more trouble is taken. The diagrams and tables arise out of a great deal of statistical transformation, some of it inherently arbitrary; and to that extent the high technical skills that have produced these results also protect them from close scrutiny by a nontechnical reader. Much is accounted for by postmaterialism, but there are also setbacks, especially in the United States and Britain. Both countries have continued to pile up affluence (albeit more slowly than before) but have also experienced an extended hegemony of the Right, intellectual as well as political, which would seem to go against the grain of the theory. Penal policy has become more punitive, the death penalty has returned in the United States, inequality has widened and the majority (especially in the United States) are not much better off. It could be argued that setbacks to postmaterialism reflect a failure of economic growth and distribution, but that is not a line that the author takes. Recently, both President Bill Clinton and Prime Minister Tony Blair appear to be choosing quite a strong materialist and traditionalist orientation;

perhaps the postmaterialist base is not firm enough for a postmaterialist leadership.

From Subject to Citizen: Australian Citizenship in the Twentieth Century. By Alastair Davidson. New York: Cambridge University Press, 1997. Pp. xiii+341. \$64.95.

Jeff Spinner-Halev
University of Nebraska

In his ponderous book, *From Subject to Citizen*, Alastair Davidson ably traces the political and legal development of Australian citizenship in the 20th century. There are three movements of citizenship that interest Davidson. The first is the move from subject to citizenship. As Australian citizenship for much of the 20th century was defined in Anglo-Celt terms, the second movement is the recognition of the multicultural nature of Australian citizenship. The third is the development toward an "active citizenship" instead of a passive one.

Davidson claims that this third development has not been successful, a failure he traces to the way Australian citizenship developed. In contrast to other Western democracies, people in Australia never fought for their citizenship rights. Since most of the original white settlers in Australia were convicts or ex-convicts, they were gradually given rights by the "grace and favor of the state" (p. 140). Over the course of the 20th century, Australians became citizens, with similar rights to those in other Western democracies, though Davidson argues that some of these rights are tenuous and should be enshrined in a bill of rights. Davidson's real concern, though, is with the passivity of Australian citizenship. Davidson contends that the modern democratic citizen established him or herself "over centuries of struggle against the state power, and has fought to retain those rights ever since" (p. 254). The British, French, Americans, and Italians, for example, all fought the state for their citizenship rights. The Australians, however, never went through a similar experience, which explains their apathy.

But are citizens in the United States or Western Europe more active than those in Australia, as Davidson assumes without evidence? American voter turnout is notoriously low, with many Americans turned off by politics. While Davidson complains that Australians have not fought hard enough for their citizenship rights, again the evidence he provides is scanty. Davidson mostly thinks of citizenship in terms of voting rights, which is surely important. But measures of active citizenship need to go beyond voting. Do Australians join interest groups? Neighborhood associations? Labor unions? Do they protest and organize? And do they do these things more or less than citizens in other Western democracies? Davidson hardly broaches these questions.

Australian citizenship was notoriously racist for much of the 20th century, denying citizenship rights to Aborigines and Asian immigrants and their descendants. Davidson argues that now, at the close of the 20th century, Australian citizenship is truly multicultural and indeed something of a model for other states. Davidson's analysis of Aborigines concentrates on their voting rights. Yet one may wonder whether simply giving Aborigines the right to vote is the same as giving them full citizenship. Davidson notes that giving Aborigines the vote will not ensure that the government will be responsive to their needs. Davidson obliquely claims that large governments tend to ignore the problems of small groups; the problem will not be solved until there is a "devolution of decision-making from the centre" (p. 199).

Davidson should have better explained this assertion. He also should have engaged Will Kymlicka's well-known argument that indigenous peoples deserve special, group rights to protect their cultural structure. Kymlicka argues that, even on a local level, indigenous peoples need special rights because they can be outvoted by whites who have settled in the area. Davidson, however, dismisses Kymlicka's argument by simply saying that such an approach "seems risky" (p. 146). It is worth noting that Davidson does not single out Kymlicka for inattention: his almost complete lack of attention to the literature on multiculturalism is striking.

The evidence of a multicultural Australian citizenship is that immigrants from all over the world can now receive Australian citizenship and that the requirements of a command of the English language for naturalization has been relaxed. Davidson suggests different cultural communities now flourish in multicultural Australia. Yet he also points out that intermarriage rates are quite high, raising the question of how distinctive these different cultural communities really are. While one may be able to now vote in Australia with minimal knowledge of English, Davidson says little about the need to speak English for social and economic equality. Can one succeed in Australian schools and businesses without knowing English well? If not, then relaxing the need to know English to become a citizen hardly means the ushering in of a new multicultural era.

The requirements for Australian citizenship are no longer racist as they once were. Readers interested in tracing this development, as well as the movement from subject to citizen in Australia, would do well to consult this book. Those who seek a deeper understanding of active or passive citizenship in a diverse Australia, however, will want to turn elsewhere.

Molding Japanese Minds: The State in Everyday Life. By Sheldon Garon. Princeton, N.J.: Princeton University Press, 1997. Pp. xvii+313. \$24.95.

Tetsuya Kataoka
Hoover Institution

Molding Japanese Minds concerns the efforts of the Japanese state to intervene in the everyday life of its citizens to harness them to national purpose. The author does not deny that modern Japan's mobilization of the populace was "not so subtle" as in the West. "The Japanese campaigns," he says, "remind one of wartime drives launched by the Anglo-American powers to maintain popular morale and promote savings and austerity during the two world wars" (p. 13). If Japan is "a nation at war in peace," then it would be an authoritarian movement regime under "the emperor system," which is what defined Japan for E. Herbert Norman and the *Kozaka* school of Marxists.

But Garon says he wants to eschew "the emperor system" paradigm or even Michel Foucault's concept of "social control." Rather, he wants to conceptualize modern Japan's popular mobilization as "social management," in his words. Social management would seem to be subtler and more refined than the imperial Special Higher Police or social control. And he intends that it is. Garon says he wants to get at what Japanese officials call *regulation* rather than the control (*tosei*) used only in wartime.

Regulation, he says, "connoted not so much the suppression of popular activity . . . rather, regulation meant that the authorities officially recognized certain activities and groups, while placing them within supervised frameworks or 'zones'" (pp. 6-7). As long as social groups stayed within the zones, they were granted a fair amount of autonomy. The ties that bound the state to daily life were "moral suasion," "social education," and "spiritual mobilization," according to Garon. In essence, he is trying to get at that softness in texture that characterizes things Japanese. As John W. Dower noted, at the height of the life-and-death struggle in World War II, the wartime propaganda of imperial Japan was less pungent and stinging than the American counterpart (*War without Mercy* [Pantheon, 1986]).

With skill and finesse, Garon has brought to light the hows and wherefores of social management over the long span of time from the Meiji Restoration (1868) to 1945, on such varied topics as the welfare system, religious policy, licensed prostitution, and women's organizations. In part 2 of the book, he extends the study beyond 1945 to the present. His attentiveness to the role of gender enables him to show for the first time the efficacy of enlisting housewives in the savings drive, among other things.

As he describes it, social management rested on consensus and voluntary cooperation between officials, primarily of Home Affairs, and the populace, of which the rising middle class after World War I acted as cadres, linking the state and society. He presents a picture of more or

less seamless linkage between government bureaucracy and civic groups. Garon's Japan, then, is a modernizer that sought to uplift the society through moral suasion rather than by acting as an "oppressor."

Molding Japanese Minds will be welcomed by Japan specialists as well as sociologists studying religion, women, suffragists, and economic development across cultural boundaries. I think this is a splendid book, but there is a strange incongruity between the title and the theoretical chapters, on the one hand, and the body of research, on the other.

As noted, Garon posits a blanket judgment at the outset that Japan is "a nation at war in peace" because of its robust intervention in everyday life to mold Japanese minds. Moreover, he thinks this judgment applies equally to prewar as well as postwar Japan. "If one looks beyond a few controversial issues . . . such as the role of the emperor or the status of the postwar military," he says, "we discover many areas of consensus." He goes on to say, "Relations between the state and society in postwar Japan underwent . . . hardly a revolution" (p. 152).

Is this judgment tenable? Did not the occupation reforms capped by the scrapping of the Meiji Constitution rate as a "revolution," as Charles Kades, MacArthur's aide, said? Karel van Wolferen, the Dutch journalist to whom Garon seems to pay rapt attention, maintains that today's Japan is a "stateless nation," a condition that can be traced directly back to the new constitution (*The Enigma of Japanese Power: People and Politics in a Stateless Nation* [Knopf, 1998]). Is he not correct? Did not licensed prostitution—a subject of Garon's study—go out of existence as a direct consequence of enfranchising women, as mandated by the MacArthur constitution?

In my judgment, yes should be the answer. What does that mean for Garon? It seems to suggest that his theoretical apparatus—a marriage of the modernization paradigm and the state-society distinction—is still evolving and remains to be sorted out.

Let me note one factual error. Hatoyama Ichiro, prime minister in 1954–56 and grandfather of Hatoyama Yukio of the Democratic Party today, was not purged because he was a "jingoist" or "pro-fascist" (p. 168). The occupation purge was an administrative procedure in which there was no trial or appeal. It sought to wipe out a social class, and individual "guilt is irrelevant," as one general headquarters official said at the time. All people purged were by definition "extreme militarists." The elder Hatoyama was given that dunce cap because he had the temerity to denounce the communists at the Foreign Correspondents' Club in 1945. He was in fact framed by Mark Gayne, a reporter of the *Chicago Sun*.

The Rule of Law in the Arab World: Courts in Egypt and the Gulf. By Nathan J. Brown. Cambridge: Cambridge University Press, 1997. Pp. xvii+258. \$59.95.

Lawrence Rosen
Princeton University

To many Westerners the Arab east may seem an inauspicious venue for studying the rule of law. But as Nathan Brown is quick to point out, resort to courts is not only popular in countries of the region—especially in Egypt—but a study of the reasons why, from colonial times to the present, the regimes of the area have supported a potentially counterbalancing institution poses indispensable questions for students of political sociology. In the course of his comparison of legal institutions in Egypt, Kuwait, Qatar, Bahrain, and the United Arab Emirates he tests the adequacy of explanations based on elite support for liberal legality, the role of litigation in the development of democratic institutions, the importance of law as a legitimizing force, and the unintended consequences that may flow from using quasi-independent institutions to bolster centralized authority.

Starting with his central case study of Egypt, Brown argues that although the British displaced the French they did not displace the influence of French legal forms but melded them with their own colonial emphasis on codes and legal centralization. Local elites did not oppose such developments, seeing the courts not as a counterweight to authority so much as a way of regulating it in order to make it more compatible with modern productivity and efficiency. Special courts for foreigners and political cases were largely merged in the unified system that was implemented in the 1930s, but once again administrative rationality rather than administrative counterweight was the driving force for reform. By the time the military came to power in 1952 and the ideology of socialism became the dominant voice of politics, legal personnel could be seen as servants of the revolution, and what independence some had shown was largely submerged in the new statism.

It is precisely because this history appears to be largely defined by law in service of the state that what next occurs is so striking. For in the last quarter century Egyptian courts have, on occasion, stood up for just that liberal legality that had never gained preeminence earlier. Brown has set the groundwork, though, by showing that this element was always present, not because of foreign imposition or slavish imitation of Western forms but because, somewhat ironically, the authoritarian regimes, whether colonial or indigenous, saw the courts as arms of the state, thus bolstering the courts' own standards for appraising authority. Thus the Supreme Constitutional Court, established in 1969, has occasionally ruled against the political powers. That special courts once again siphoned off some of the more troublesome issues has actually contributed to the independence of the courts in other domains; that unification of the courts

was once in service to socialist ends and is now conceived in more liberal terms is not a tribute to democracy, however, but to the coincidence of judicial and political interests, such that even the occasional judicial restriction on executive action is seen as the furtherance of a common statist goal. Brown's thesis, in effect, is that law serves not a legitimizing role in the eyes of Egyptians but that the instrumentalism of all members of the polity, which might seem to tie the hands of the executive at certain moments, is sufficiently restrained that it can occasionally play a supportive role that is, ironically, also not unlike a partial limitation.

In the Gulf states, although Islamic courts continued to influence developments to varying degrees, the Egyptian model was largely applied. As Brown says: "Gulf rulers have been drawn to legal reform [because] law codes and codified law increase the authority and effectiveness of the central state" (p. 185). Political hierarchies have been less challenged than in Egypt, thus suggesting that the extent to which judiciaries can foster liberal legality is a function of the larger set of political institutions.

Indeed if there is a limitation to Brown's explanation for why Egyptian and Gulf officials have supported degrees of judicial autonomy it lies in his partial consideration of the larger cultural and political context. There is more to legal culture than elite politics. Islamic law, for example, is not only about matters of personal status: it also reflects an ethos of limited central power, as embraced in procedures, assumptions, and forms of reasoning that reinforce similar attitudes in other domains. It may be that Egypt was, unlike Iraq, never authoritarian enough to eradicate the ethos of that larger culture which, in humor, drama, literature, and law, has helped to maintain a degree of limited central authority. By adding these considerations to this carefully constructed and thoughtful work, the deeper meaning of the rule of law becomes even more available for comparative study.

Jews in Germany after the Holocaust: Memory, Identity, and Jewish-German Relations. By Lynn Rapaport. New York: Cambridge University Press, 1997. Pp. xi+325. \$59.95 (cloth); \$22.95 (paper).

Andrei S. Markovits
University of California, Santa Cruz

As if to confirm the continued special nature of German-Jewish relations—in case any were still needed—the *Boston Globe* ran a front-page article entitled "In Berlin, Jewish Is In" on December 23, 1997, the very day I concluded reading Lynn Rapaport's fine study. The article informed its readers that in no other European city can one find such a proliferation of klezmer bands as in contemporary Berlin; nowhere else do people stand in line for three hours to enroll in Hebrew, Yiddish, and Old Testament courses; and few, if any, cities can boast of a Jewish choir as accomplished and versatile as the one in Berlin. Rendering this sce-

nario particularly fascinating is the fact that none of the participants in any of these ventures is Jewish. The performers are Germans, as is the audience; the students as much as their teachers. The *Boston Globe* article's subtitle sheds more light on this peculiar phenomenon: "Germans Study Culture Nazis Tried to Destroy." And they do so with a vengeance it seems.

To anyone familiar with the complex phenomenon of German-Jewish relations since the Holocaust, the scenario described by the newspaper article comes as no surprise. Indeed, it corroborates once again the uniquely twisted nature of this relationship that finds its most erudite exposé and most serious scholarly analysis to date in Lynn Rapaport's book. The argument is simple: As a consequence of the Holocaust and the collective memory of it, no interaction between Germans and Jews—no matter how fleeting or substantial—exists outside the prism of this horrible event that continues to shape the lives of Germans and Jews in the modern industrial democracy better known as the Federal Republic of Germany. Everything that each of these two groups does is perceived by the other through the prism of the Holocaust. Thus, the result is one of constant distortions, exaggerations, excessive sensitivities, misunderstandings, and overreactions, which in turn create their very own new realities. As Rapaport amply describes, the Jews perceive most of these realities as negative. While she explicitly does not deal with the German side of the story (one of the book's shortcomings, in my opinion), it is obvious that for the Germans, too, things are far from straightforward—or "normal"—in their relationship with Jews. Among the many reactions on the part of the Germans has been an overt philo-Semitism and an exaggerated interest in things Jewish, as evidenced by the Berlin story. Telling of the immense complexity of the German-Jewish relations, this philo-Semitism on the part of the Germans has far from made the Jews comfortable in their feelings toward Germans and Germany. It is this discomfort that informs the empirical substance of Rapaport's study.

Through in-depth interviews with 83 members of the Frankfurt Jewish community conducted repeatedly in the decade between 1984 and 1994, Rapaport conveys powerfully how this group continues to feel marginal, alienated, afraid, unintegrated, reviled, uncomfortable, yet also admired and appreciated but most certainly different from its German surroundings. And the Jews do so—according to Rapaport's main thesis—by dint of one single matter: their (and the Germans') collective memory of the Holocaust. Indeed, it is this memory clash that continues to differentiate Jews from Germans in the contemporary Federal Republic and that furnishes far and away the most potent ingredient in giving Jews their Jewish identity in today's Germany. Rapaport shows nicely how today's Jews in Germany—who, tellingly, would never identify themselves as German Jews—possess all the characteristics of being an indistinguishable segment of German culture and society. Germany's Jews carry German passports, they speak German as their native tongue, they are completely conversant with German culture, they are solidly middle class, they enjoy

the material comforts of one of the world's richest societies, they have full political rights, they are employed in the most respected professions, and they can come and go as they please. Moreover, as a rule, they do not speak Hebrew; they are almost completely secular, confining their religiosity to three synagogue visits per year for the High Holy Days in the fall when they attend synagogue services mainly for social rather than religious reasons; they do not keep kosher; they do not observe the Sabbath; and they do not speak Yiddish. In short, they would be indistinguishable from the Germans but for one major fact: their collective memory about the Holocaust. This, as Rapaport demonstrates in impressive detail, affects all their perceptions of and interactions with Germans: on the job, in private life, as coworkers, friends, and lovers. Abstractly and concretely, in theory and practice, for Jews in contemporary Germany, the Holocaust is *always* present.

Rapaport's fine study demonstrates powerfully how a deeply experienced "otherness" can reproduce itself daily in a situation devoid of any objective differences and ostensible discrimination. Indeed, it can—and will—do so in a context where the former "outgroup" has attained a position of informal privilege and even admiration from the rest of society, as attested by the *Boston Globe* article as well as examples from Rapaport's book. But there is simply much more to group identity than objective measures and measurable criteria. One almost has the feeling after having read Rapaport's excellent book that were it not for the Jews' collective memory about the Holocaust, the Germans would finally be ready to accept the Jews as an integral part of their society and culture. This, alas, is an erroneous view that Rapaport could have countered very easily by devoting at least one chapter to the Germans' attitudes and behavior toward the Jews. While it is clear that in the pre-Holocaust world many of Germany's Jews were desperate to have the Germans accept them as Germans (of the Jewish faith if need be), it is obvious that Auschwitz would only permit Germany's contemporary Jews the much looser identity of Jews currently residing in Germany. The Germans did not want the Jews to be Germans before 1945; now, the Jews do not want to be Germans as Rapaport shows. As to whether the Germans want the Jews as Germans in the contemporary Federal Republic still remains an open question in this reviewer's skeptical mind.

Internationalism and Its Betrayal. By Micheline R. Ishay. Minneapolis: University of Minnesota Press, 1995. Pp. xlviii+180. \$17.95 (paper).

Roman Szporluk
Harvard University

This is a short book with a very wide thematic range. In part 1, "The Disintegration of the Universal Church and the Development of New World Vision," Micheline R. Ishay reviews the ideas of Grotius, Vico, and Rousseau. This is followed by a discussion of Kant, Paine, and Robespierre under the common theme of "Internationalism and the French Revolution." In "The Rise of Nationalism and the Counter-Enlightenment," the author addresses the ideas put forth by Burke, de Maistre, and Fichte. The book concludes with an essay titled "The Great Fusion: A Commentary on Hegel's *Philosophy of Right*." The foreword by Craig Calhoun is a rare example of an effective introduction to what follows.

Ishay argues that internationalism preceded nationalism and that it needs to be distinguished from cosmopolitanism and universalism, both of which had come earlier than internationalism. She offers various descriptions, if not exactly definitions, of internationalism but perhaps the following gives a clear indication of what she has in mind: "Internationalism has to be understood . . . as philosophical guidelines describing social relations between and within states" (p. xxxi). Ishay shows that the traditional debates in political thought about good government, the relations between citizens and their rulers, the liberties and duties of citizens, and so forth also had an international dimension. (International in the sense she understands the word and as it was understood in the time of Grotius).

Internationalism emerged, Ishay argues, concurrently with the decline of Catholic universalism, the Reformation, and mercantilism. As her discussion of Grotius makes clear, she uses *nation* basically as another word for the state that emerged in Western Europe as the universalist Holy Roman Empire declined. We used to call those West European states *nation-states*. Thus it is obvious that she uses *state* and *nation* interchangeably. This usage is of course not uncommon: we speak about *international law* or the *law of nations* when everybody knows that "international" law regulates, or is supposed to regulate, relations between *states*. Since this terminology is so out of fashion in the literature on nationalism, one wonders if she would not have made her position clearer if she spoke of "inter-nation-stateism" or "inter-nationism" instead of speaking of internationalism.

These are not just terminological quibbles. This book develops a point of view that is very different from what is widely accepted in contemporary scholarship on nations and nationalism. Few authors writing on nationalism today would agree with Ishay that before nationalism there was internationalism—but the disagreement has its roots in a very different

understanding of what nation is. Thus many would question her claim that after 1815 in countries such as Britain, France, and Spain "national unity had already been achieved" and that nationalism in those countries "was not so much a doctrine as a latent state of mind, easily aroused when 'national interests' were questioned" (p. xlv). They would respond by citing books about how "peasants" were turned "into Frenchmen," about the Irish problem and British identity, about Catalonia, the Basques, and other problems.

To say that this book takes a line very different from current fashions is not to question the author's right to offer her alternative view of nation, state, internationalism, and nationalism. Ishay advances an original argument that deserves a hearing. But she would have made a stronger case if she had engaged in a direct dialogue with such influential writers on nationalism as Benedict Anderson, Ernest Gellner, Liah Greenfeld, or Eric Hobsbawm. While these scholars deal primarily with a more recent period of history than hers, there is in fact a good deal of chronological and thematic overlap between her book and those of Benedict Anderson (*Imagined Communities: Reflection on the Origin and Spread of Nationalism* [Verso, 1983]) and Liah Greenfeld (*Nationalism: Five Roads to Modernity* [Harvard University Press, 1992]) in particular. Similarly, judging by Ishay's footnote references, her views of thinkers she is studying are based on her own reading of their original texts. But it would be good to hear from the author where her views differ, when they do, from current interpretations of such figures as Vico, Rousseau, or Hegel in the scholarship on these thinkers. This too would make her message more interesting to a wider circle of readers.

Social Darwinism in European and American Thought, 1860–1945: Nature as Model and Nature as Threat. By Mike Hawkins. New York: Cambridge University Press, 1997. Pp. x+344. \$69.95.

Leonard Moore
McGill University

In the last decade of this century, as scientists have raced to map the human genome, as experimental animals have been cloned, and as some scholars have resurrected disturbing questions about the inherent intelligence of different racial groups, controversy has surrounded the social implications of genetically determined human characteristics. This, of course, is nothing new. From the time of Charles Darwin's *On the Origins of Species*, to the eugenics, birth control, and sterilization movements at the turn of the century, to the laws against teaching evolution of the 1920s, to the discovery of DNA in the 1950s, to contemporary debates about sociobiology and the ethics of new reproduction technologies, contending ideas about the meaning of human genetics have inspired deep, passionate disagreements.

Social Darwinism, as most people understand it, symbolizes the worst elements of this contentious history. Embraced as an axiom through the early 20th century, then just as widely repudiated after World War II, social Darwinism is popularly perceived as a rationalization for racism, imperialism, patriarchy, and class exploitation.

Mike Hawkins's fascinating intellectual history of social Darwinism argues that this perception, while certainly justified on many levels, is also simplistic and misleading on others. By clearly defining the components of what he calls a social Darwinist worldview, Hawkins is able to demonstrate that this deeply ingrained set of assumptions "provided the philosophical infrastructure" of a wide and complex range of distinct, sometimes conflicting, ideological perspectives in Western thought from 1860 until the end of World War II. Herbert Spencer, William Graham Sumner, and others may have developed harsh social theories based on a social Darwinist perspective, but social Darwinism, itself, Hawkins argues, "cannot be defined by its ideological functions or equated with a particular political position" (p. 296).

The social Darwinist worldview, Hawkins asserts, consists of five basic assumptions: that organic nature, including humans, is governed by biological laws; that the pressure of population growth on resources brings a struggle for existence; that advantageous physical and mental traits in this struggle can be spread through a population by inheritance; that the cumulative effects of selection and inheritance cause new species to emerge and others to disappear; and that the forces that determine physical existence also apply to human social existence. Of these five elements, the last is most critical, the one that converts a Darwinist into a social Darwinist.

Darwin himself, from this perspective, was not only a social Darwinist, but "one of the major architects of the world view." From the beginning, Darwin "sought to apply evolutionary theory to mental and social phenomena" such as language, emotional expression, cognition, sexual attraction, and morals. While never developing a formal social theory or ideology, he did believe, among other things, that natural selection had elevated civilized Anglo-Saxon nations over "savage" ones and men, intellectually, over women. And yet Darwin was also a passionate critic both of slavery and of Spanish and Portuguese policies toward native peoples in Latin America. Even Darwin's social Darwinist ideas, therefore, were contradictory and complex.

The bulk of the book is devoted to a comprehensive and convincing demonstration of the fact that social Darwinism could produce all shapes and sizes of ideological expression. Early social Darwinists, such as the British journalist Walter Bagehot and the Italian anthropologist Cesare Lombroso, used human evolution to explain the need for restraining primitive social influences, respectively, in the areas of politics and criminal justice. Yet, the American linguist Charles Brace used Darwinist ideas to justify his Christian humanism (he did not believe in drawing important distinctions between races since all humans were of the same

species), and the German philosopher Ludwig Buchner used them to support his radical views on materialism and social equality. Herbert Spencer greatly popularized and provided scientific authenticity for the view that Western culture had evolved beyond the inherent savage violence, immorality, and low intelligence of primitive societies. Spencer, himself, however, struggled to reconcile this with his own findings (similar to Ruth Benedict's) about harmony and equality in some hunter-gatherer societies and his deep belief in laissez-faire capitalist liberalism. Social Darwinism produced myriad conflicting eugenics theories. German Nazis were social Darwinists; Italian fascists were not. Modern sociobiologists definitely are social Darwinists, and their claims to the contrary reflect an ignorance both of the concept and its history.

This book deserves high praise for the clarity it brings to a complex subject and for the exceptionally wide scope of its analysis. Anyone interested in social Darwinism or the many controversies that have surrounded genetics and society will benefit from reading it. It would seem particularly useful for undergraduates being introduced to the history of these topics. This book certainly will not achieve the seemingly impossible task of resolving modern nature-nurture controversies. Given his rendering of complexity and ambiguity in the history of social Darwinism as a worldview, however, one does wish Hawkins had been more enlightening about sociobiologists' claims that we can have the science without the hierarchy. He generally skirts the issue in favor of demonstrating that sociobiologists are social Darwinists as he defines them. This is not a major criticism. Clarity, definition, and scope are significant enough contributions to this controversial field.

Emancipation and Illusion: Rationality and Gender in Habermas's Theory of Modernity. By Marie Fleming. University Park: Pennsylvania State University Press, 1997. Pp. 243. \$40.00 (cloth); \$17.95 (paper).

Jodi Dean
Hobart and William Smith Colleges

Marie Fleming challenges Habermasian universalism for its failure to elaborate a vision of gender equity. To this end, she uses feminism to confront Jürgen Habermas's conception of rationality. There, she argues, rest the deep-seated gender prejudices that render his theory useless for feminists. In short, Habermas is not universalist enough.

This book is primarily a series of highly detailed reconstructions and engagements with various articles, books, debates, and interpretations by and of Habermas, rather than a sustained discussion of universalism or gender. The author's own suggestions for a reconstructed universalism are hinted at on the last page. Half the chapters do not deal with gender at all.

In the first section, Fleming introduces the concept of communication

action and the principle of universalization at work in discourse ethics. She faults the former for its Eurocentricism and the latter for its "selectivity." "How is it that some values, those related to 'good life' issues, get classified as 'particular values [that] are ultimately discarded as being not susceptible to consensus?'" (p. 80).

This is an odd question. The answer to "how" is relatively straightforward: in the course of a practical discourse. Habermas does not determine in advance what goes under the headings "justice/morality" and "ethics/the good life." Rather, he defines the categories in terms of universalizability. For Habermas, to say that a norm is "just" is to say that it *has met* with the agreement of all concerned in a practical discourse. One may well disagree with this definition. But this is not where Fleming targets her critique. Her problem is with "selectivity." At stake is her worry that issues of gender are excluded from the moral domain, relegated to "secondary" questions of the good life. This has long been an issue for feminists. Habermas's distinction between justice and the good life, however, does not require that gender be treated as ethical rather than moral. Questions of inclusion, reciprocity, equality, and recognition all challenge gender-based restrictions. Moreover, the most interesting questions may well be those ethical and political issues that are not universalizable. Indeed, the particular problems of historically situated peoples are more in need of reflective attention than the abstract concerns of idealized participants in a discourse that can never actually take place.

In the book's second section, Fleming turns to gender. Key issues are Habermas's system/life world distinction and his thesis of the juridification of the life world by the state. Fleming faults Habermas for a latent traditionalism on gender issues as well as for not providing "advice on how to deal with the issue of basic rights for women and children" (p. 92). The argument on traditionalism would be stronger had Fleming not repeated Habermas's heterosexual bias and had she drawn attention to the racism of the nuclear family model. Her discussion of juridification and Habermas's silence on basic rights is vague: she does not specify what is included under basic rights nor say why women and children should be linked together here. Other questions arise as well: Why does she assume that feminists want equality? Equality with whom and in what regard? Why does she expect Habermas to provide feminists with advice? And, more problematically, why does she not engage the voluminous writings of feminists on questions of rights, law, equality, and the family? Fleming also faults Habermas for being too general to address the question of gender helpfully. But why not read feminists for this? The effect of Fleming's reluctance to engage feminist debates is the recentralization of precisely that thinker she wishes to critique.

The most interesting analysis in the book involves Habermas's brief discussion of suttee, the Hindu ritual burning of widows, in a 1984 interview. Fleming uses the example to investigate larger patterns relating to gender, culture, and rationality in Habermas's work. Unfortunately, this analysis suffers from the absence of any links to any of the following:

postcolonialist critical theory, the debates over the "constitutive" nature of particular practices for a culture, and the writings by Indian feminists on suttee. Fleming limits her critique to the question of why Habermas fails to consider the topic in terms of rationality, a failure she sees as preventing the theorist from engaging "participants of the Hindu culture in the context of an understanding of rationality that 'they' implicitly share with 'us'" (p. 220).

A final complaint Fleming raises is that the concept of the life world is flawed because it does not provide a vision of gender equity. This contention makes sense when understood most generally: Habermas does not see any power relations as structuring the life world. To the extent that gender is a vehicle of power and is produced within relations of power, this omission makes sense. Oddly, the failure of Habermas to provide a vision of gender equality in the life world suggests an acknowledgment of power rather than its denial, for gender relations in the life world are certainly not equitable.

The New Public: Professional Communication and the Means of Social Influence. By Leon H. Mayhew. New York: Cambridge University Press, 1997. Pp. xi+332. \$59.95 (cloth); \$22.95 (paper).

Nicholas Garnham
University of Westminster

In recent years the concept of the public sphere has acted as one of the major focusing, interdisciplinary concerns of critical inquiry in political science, social theory, and media studies. It is to this general field of inquiry and debate that this book makes its impressive contribution with a fruitful and cogent development of Habermas's work on the public sphere and communicative rationality.

Mayhew starts with the same definition of the problem as Habermas—namely, whether it is possible and if so how, given the social differentiation of modernity, to realize the Enlightenment goal of a social order governed not by coercion but by the free and rational agreement of its participating subjects. His view of the present state of liberal democracy is also remarkably similar. His new public is, in effect, Habermas's refunctionalized public sphere. It designates a process of "the rationalization of persuasion" whereby professional communicators, using techniques of public and public opinion formation and manipulation developed in advertising and market research, have come to dominate the sphere of public communication and the political process based upon it using a "rhetoric of presentation." And this leads him to argue that "the single most important question brought to the fore by the arrival of the New Public asks whether there are, in the public arena, forums that institutionalize reasonable rhetoric in the face of manipulative efforts of professional persuaders" (p. 43). This book is an attempt to answer this question.

This way of posing the question indicates immediately the ways in which Mayhew attempts to develop beyond what are widely seen as the idealist impasses of Habermas's position. First, and most important, he moves from a position based upon the hyper rationality of communicative rationality to a recognition of rhetorical persuasion as central to any system of social action based upon consent. Drawing on classical rhetorical theory and the work of Kenneth Burke, he argues that, while we can think of rhetoric as "blending, in persuasive action, reason and four other elements of effective communication: norms, solidarity, coercion and eloquence" (p. 38), this does not mean that we cannot base a public sphere on "reasonable rhetoric," that is, rhetoric that makes validity claims that can be redeemed. For Mayhew the claim to social solidarity is of particular significance.

Second, he returns to one of the key roots of Habermas's theory in the work of Talcott Parsons to argue, against Habermas, that we can see influence, defined as a generalized capacity to persuade, as a medium in Parsons's sense of the term. Here he makes a realist move away from Habermas, drawing both on Parsons and on work in the economics of information, to argue that people have neither the time nor the energy to participate in the fully rational discursive practice required by Habermas's public sphere and thus the sphere of influence necessarily uses as its currency "rhetorical tokens" and requires representatives (what he calls *prolocutors*) to speak on others' behalf. As with any medium the circulating currency needs to be underpinned by a system of credit, that in the sphere of influence is reputation. Here Mayhew makes a crucial break with Parsons, who made the existing social hierarchy the foundation of reputation. For Mayhew reputation is a token that is itself always in principle redeemable within the sphere of influence. And while the rhetorical claims of a prolocutor always appeal, he argues, to a social solidarity, this solidarity may not be preexistent but may be called into existence by rhetoric itself. Nonetheless the reputation and thus influence of all prolocutors depends upon the validation of the rhetorical claims made through endorsement. As he puts it, "In modern society influence becomes differentiated from the general status order as it comes to be based on resources that must be created by rhetorical bids to speak for particular, often new or emergent groups and alliances of people who can be brought together under a banner of common interest" (p. 126).

From this general position Mayhew concludes by arguing that "only institutions requiring those who exercise public influence to enter forums wherein their token claims can be redeemed in information and argument could secure the integrative potential of influence" (p. 110) and that the problem with the new public is "the growing separation from the social organization of solidarity that makes problematic the methods of persuasion utilized in the New Public, techniques designed to appeal to manipulable individual attitudes rather than to attach citizens to likely allies and to collective support for programmes of public action" (p. 184). He then ends the book by suggesting ways in which what he calls "forums for

the redemption of influence" might be constructed by building upon the experience of the civic journalism movement, deliberative polling, and the National Issues Forum. This, not surprisingly, given the difficulties involved, is the weakest section of the book, but Mayhew has provided us with a sound basis for this necessary work of political imagination and institution building by providing a cogently argued realistic development of Habermas's theory that meets the major current objections to it.

World in Fragments: Writings on Politics, Society, Psychoanalysis, and the Imagination. By Cornelius Castoriadis. Edited and translated by David Ames Curtis. Stanford, Calif.: Stanford University Press, 1997. Pp. xi+507. \$65.00 (cloth); \$22.95 (paper).

Thomas M. Kemple
University of British Columbia

We have become accustomed in recent years to treating the fragmentation of our knowledge of the world as itself a symptom of our current world crisis. This has led a chorus of critics to call for the institution of a truly public sphere that would acknowledge the stratified multiplicity of our social and personal existence (including its racist, conformist, and even psychotic character) by rethinking our *psychic and social imaginary*. This is the argument that might initially attract English-speaking social scientists to the translation of Cornelius Castoriadis's latest book, particularly its first part titled "Koinonia" (communion, association), where it is surprising to read that "the French ideology" (particularly that of Louis Althusser, Jacques Lacan, and Michel Foucault) should be denounced for framing the current discussion in apolitical and insufficiently analytical terms. *World in Fragments*, itself a disparate collection of 19 essays, lectures, speeches, book chapters, and interviews produced from 1978 to 1993, continues the distinctive philosophical critique of Marxist orthodoxy, religious dogmatism, scientific objectivism, and psychoanalytic subjectivism that Castoriadis began over 50 years ago. As editor and translator, David Ames Curtis notes in his useful introduction, here Castoriadis produces an extraordinarily innovative blend of English, German, French, and Greek modes of expression (the last Castoriadis's native tongue) in ways that blur their technical and ordinary usages. Like the Cornelius of Acts of the Apostles whose dinner guests were blessed with the gift of tongues (pp. 20–22), Cornelius Castoriadis invites us to rediscover the language of the Western imagination by revisiting its first emergence in Ancient Greece, its resurgence in the European Middle Ages, and its revolutionary outbreak in the modern age.

Castoriadis is better at patiently interrogating his theoretical sources than he is at engaging in ideological polemics with his immediate predecessors or contemporaries. For this reason, the essay on "The Discovery of the Imagination" (from the section titled "Logos," which makes up half

the book) can be considered the centerpiece of the collection. Against the tendency in Aristotle (and later in Kant, Hegel, and Heidegger) to contain the imagination in a rationalized concept of mere sensation, intuition, or ideology, Castoriadis revives Aristotle's radical notion in *De Anima* (III. 7-9) that the soul never thinks without phantasm. In spite of the speculative and tentative character of this and the following chapters, the keynote of the entire work is sounded in this thesis: to use his own idiosyncratic terminology, "the ensemblistic-identitary logic" (or set-theoretical conception) of multiple, stratified, and discrete domains of existence (*phusis*) is "instaured" through the self-legislation (*auto-nomos*) of social-historical significations created by "the poietic imagination" (p. 262).

It is precisely this creative dimension of the human psychic process that Castoriadis wishes to recover in the programmatic essay, "The State of the Subject Today," and in the rest of the section titled "Psyche." Taking his point of departure from the idea introduced in Sigmund Freud's *The Project for a Scientific Psychology* (1895) that the imaginative work of the psyche begins in shock (*Anstoss*), Castoriadis formulates the distinctiveness of human subjectivity in the leaning (*éclayage*) of sublimated cultural representations on organic processes. However, exactly why this contradicts rather than complements Lacan's notion in "The Mirror Stage" of the child's imaginary experience of the "body-in-pieces" (*corps morcelé*, p. 52), or how it advances Herbert Marcuse's radical (and neglected) conception of the subversive and ethical play of phantasy in *Eros and Civilization* (Beacon, 1955), is never made clear.

Most of these writings were taken from two later volumes in the series whose first installment was translated as *Crossroads in the Labyrinth* (Harvester, 1984). The essay that gives *World in Fragments* its title in fact appears in the earlier work (although there it is called "Modern Science and Philosophical Interrogation," as Curtis notes), and a fourth section titled "Polis" appears only in the later work. Perhaps this signifies a renewed concern for issues that Castoriadis has addressed since his days with the journal *Socialisme ou Barbarie*. Nevertheless, in spite of some trenchant political commentary here in "The Ethicists' New Clothes," and in the essays on the unfulfilled promises of 1789, February 1917, and May 1968, he seems incapable of finding the lessons he draws from Aristotle and Freud in the work of Karl Marx and Maurice Merleau-Ponty (as might have been expected from the earlier volume). Marx's supposed "scientific laws" are quickly dismissed as a quest for "secular salvation," while Merleau-Ponty's exploration of the social-historical imagination is exhaustively treated only for its apparent "ontological bias." Although Castoriadis's own conceptual scheme is further clarified in these ruminations, any notion of the *political imaginary* of a heterodox Marxist theory that might inform new possibilities of human existence (p. 57) remains to be amplified in the work to come.

Back to Reality? Social Experience and Cultural Studies. Edited by Angela McRobbie. New York: Manchester University Press, 1997. Pp. vii+216. \$74.00 (cloth); \$24.95 (paper).

Herman Sanders Gray
University of California, Santa Cruz

Cultural studies has been the source of considerable misunderstanding and, on occasion, acrimonious debate. And yet, in many professional and scholarly quarters cultural studies is hot; publisher's lists, bookstores, course syllabi, and conference sessions regularly include sections that appear under its heading. Whatever one thinks about cultural studies, one thing is unmistakable—it is more than a passing fancy or the latest intellectual fashion. With an international reach, a definable body of work, and an identifiable network of practitioners and journals, cultural studies seems here to stay. Despite this success, Angela McRobbie and some of the contributors to *Back to Reality* believe that cultural studies is at a crossroads. Cultural studies needs a pause to take stock, to distinguish itself from its worst imitations, and to confront both its own intellectual practice and the rapidly changing world that it purports to examine and aims to transform.

Given the critical probity of her prior work on feminism, popular culture, and postmodernism, McRobbie is especially alert to the challenges facing cultural studies. A central theme connects and animates the book's contributions—each writer demonstrates how cultural studies analyses might proceed and challenges cultural studies to explicitly account for its own practice.

The four chapters in part 1 appear under the subtitle, "Cultural Studies in an International Frame." Lawrence Grossberg suggests that the central categories of modernism—nation, subject, agency—are no longer productive in accounting for a social and cultural world in the throes of the global flows of people, capital, information, and culture. Grossberg proposes replacing the primacy of 18th- and 19th-century conceptions of temporality and geography with what he calls spatial and territorial conception or machines. The gain is an intellectual practice that can more adroitly apprehend the new relations and sites of power and thus serve as a more productive basis for politics. Similarly, Megan Morris's challenge is to elucidate the contours of cultural studies in the context of both its intellectual practice and its politics. Popular and academic (mis)conceptions of cultural studies in Australia provide Morris with the context to elaborate and clarify in a programmatic way the cultural studies project.

The chapter by Graham Murdock will be of particular interest to sociologists since he takes up a number of questions that lie at the heart of sociological skepticism about cultural studies; these include questions of method, data, and the claims on culture and power. Murdock explicitly challenges cultural studies scholars to grapple more directly with the issue

of political economy, by accounting for the specific way in which culture shapes the material organization and production of cultural products and practices.

bell hooks contributes the only explicit discussions of pedagogy by describing the demands and rewards of a cultural studies pedagogy that aims to transform the hierarchically structured classroom into a safe and challenging space. She argues that teachers must recognize and engage the whole student, a goal that can best be achieved if we transform our classrooms into spaces where loving and passionate intellectual relationships with our students are welcome. Bold in its assertions and risky in its revelations, the chapter is part confession, part testimony, and part ethnography.

Chapters by Paul Gilroy, David Laing, Vron Ware, Maria Pini, Sean Nixon, and Angela McRobbie appear under the rubric "bodies, images, and music." In these chapters the body (Gilroy, McRobbie, and Ware), cultural practices (Laing, Pini) and social spaces (all) take center stage. By detailing the practices and spaces—the black public sphere, sex tourism, raves, retail design, and popular magazines—in which black, white, male, female, consuming, and dancing bodies operate, the authors interrogate the social relations of power, exposing the structuring forces that operate on the body through various sites of popular practice. Laing's essay takes exception to some of the ways in which cultural studies scholars have theorized the cultural significance of popular music. Essays by Nixon, Pini, and McRobbie demonstrate that cultural spaces—in this case, retail design, raves, and popular magazines—involve complex relations and codes through which masculinity and femininity is produced. Gilroy makes a similar point through his nuanced analysis of the cultural impact on the black public sphere of the loss/transformation of the consciousness of freedom in black popular culture.

Although these studies represent some of the most interesting work in cultural studies, some readers may find the volume uneven since some of the chapters assembled here have appeared before in journals or other collections. This occasionally gives the book the feeling of being dated if one has encountered them before (e.g., chapters by Grossberg, Gilroy, and hooks). On the other hand, the chapters by Nixon, Pini, and Morris appear to American readers for the first time. Similarly, for all of the cultural studies talk about globalization, the chapters focus primarily on Britain, Australia, and the United States. While hooks, Gilroy, and Ware at least make forays into the spaces where dark bodies reside and play within the West, the globalization process is confined to a few industrialized Western countries.

Back to Reality does effectively pull together some of the most significant reflections on the state of cultural studies as an intellectual formation and practice. For sociologists, it is a useful point of entry and one that offers challenging insights and rich examples of cultural studies scholarship.

Anthropological Demography: Toward a New Synthesis. Edited by David I. Kertzer and Tom Fricke. Chicago: University of Chicago Press, 1997. Pp. ix+294. \$46.50 (cloth); \$18.95 (paper).

Alaka M. Basu
Cornell University

I begin my review of this generally excellent book with a not-so-minor quibble. The editors, and sometimes the authors, are never quite clear about whether they are doing anthropological demography or demographic anthropology. This would be all right if they said that they did one or the other, but they seem to use the two terms interchangeably. The editors in particular continuously refer to the subject as anthropological demography and to its practitioners as demographic anthropologists.

There is more than semantics at stake here. To my mind anthropological demography and demographic anthropology are not merely two ways of referring to the same subject. They imply separate things about the basic training of the researcher, the methods of analysis, and some of the key variables of interest. In particular, anthropological demography implies the use of the methodological, empirical, and theoretical tools of anthropology to further demographic understanding, while demographic anthropology implies the use of demographic methods and information to throw more light on some of the key questions in anthropology, many of which are often one step removed from the fertility, mortality, and migration issues that so engage mainstream demographers.

In such a formulation, the contents of the book tilt it in the direction of anthropological demography in that it focuses on what demography and demographers can learn from anthropology. But given that most of the authors are primarily anthropologists, I wish there had also been some more introspection on what anthropology can learn from demography. For here we have the anthropologist as teacher at her best but virtually nothing of the anthropologist as student, the concluding paragraphs in the editorial introduction notwithstanding.

That said, as a demographer myself, I learned much from this volume. While I refreshed and broadened my understanding of some things that demographers have now been aware of for some years—the relevance of family and kinship structures for demographic outcomes for example—I also appreciated more keenly the pitfalls in our eagerness to translate all our statistical findings into simple-minded interventionist policy. That is, we seldom stop to ask if we should play God in people's personal lives; moreover we assume too easily that we *can* play God—population history is littered with examples of failed policy initiatives.

As the authors scold, this failure is often caused by our efforts to change behavior well before we can even half understand it. That such understanding can be particularly thwarted by the demographer's prior conceptions and definitions is the complaint of many of the articles in this volume. The emphasis on narrowly biological definitions of births and

deaths, fertility and health; the classification of people into categories that suit the researcher more than the subject (Kreager's reminder that collective identities are multiple and mutable is superb); the equation of empirical credibility with sample size and statistical sophistication; the tendency to reduce complex questions of gender to simplistic (and measurable—that is the key hallmark of the traditional demographer—everything must be quantitatively measurable) notions of education and employment are all rightly deplored in individual articles. The notion of culture in demography gets a particularly baleful eye—even the famous Princeton European Fertility Decline Project, which raised culture to more than the mere background variable in cross-cultural comparative demographic analyses, never really led to a grappling with the complicated and dynamic nature of culture and the continuing importance of agency in cultural explanations of behavior.

The book also clarifies for demographers that anthropological demography is much more than the adoption of the field procedures of anthropology. In addition, it is much more than just a microperspective. The whole emphasis on "context" that characterizes the anthropological approach cannot but include attention to the larger institutional, historical, economic, and political structures that interact with individual behavior as much as do cultural norms.

My only dissatisfaction with this book is a pragmatic concern. *Who* is the typical anthropological demographer? Is it one person with two hats? Or is it a team of trained demographers and anthropologists complementing one another? In the former case, we need to think more about cross-disciplinary training. In the latter, the onus is on mainstream anthropology, which is so accustomed to working alone and is usually so possessive about its territory.

Auto/Ethnography: Rewriting the Self and the Social. Edited by Deborah Reed-Danahay. New York: Berg, 1997. Pp. xi+277. \$55.00 (cloth); \$19.50 (paper).

After Writing Culture: Epistemology and Praxis in Contemporary Anthropology. ASA Monographs 34. Edited by Allison James, Jenny Hockey, and Andrew Dawson. New York: Routledge, 1997. Pp. ix+273. \$19.95 (paper).

George E. Marcus
Rice University

Taken together, *Auto/Ethnography* and *After Writing Culture* offer a very good survey of what has been made of the 1980s critiques that have effectively marked the remaking of how social and cultural anthropology in the United States is written and, to some degree, practiced. They also reveal something of the disciplinary politics that have been at stake in

this change. Of the two, *Auto/Ethnography* is the more thematically focused and specified. It looks at how the genre of life history in ethnography has become a primary mode of experimentation. In her excellent introduction, Deborah Reed-Danahay points out that the central thread of the volume is its concern with life stories—"sometimes our own, sometimes those of others, and sometimes both. The act of self-narrative and the tension between creativity and restraint associated with that act in various political contexts, is a central theme in this book" (p. 1). The papers are defined by their overlapping and mixing of three genres: "native anthropology," "ethnic autobiography," and "autobiographical ethnography."

Several of the papers in *After Writing Culture* address similar questions as those in *Auto/Ethnography*, but its concerns lie more with the problems of representational practices in ethnography and therefore with various responses to the 1980s critiques that prominently define current disciplinary politics. The most provocative pieces in this volume display strong feelings, one way or the other, about the radical implications of the critique: for example, the focal concern with positioning, out of feminist theories of the self, as defining the central problematic of contemporary ethnography, addressed in Lisette Josephides's useful review of three influential works by Anna Tsing, Ruth Behar, and Kemala Visweswaran; Glenn Bowman's examination of the cultural (even deterministic) embeddedness of the way that anthropology defines the self that creates ethnographic representations, the suspicion, even resentment, in Declan Quigley's paper about the way that attacks on representation have derailed scholarly traditions, as in the literature on Indian caste; and the way that the problem of representation matters within applied anthropology in Sandra Wallman's paper.

What seems to me implicit in many of these papers is an inevitable predicament following upon the radical critique that challenged the very "doing" of anthropology—that of defining one's position defensively so as to stay away from blind alleys that would lead to paralysis. In neither of these volumes does anyone bemoan the "end of anthropology," but neither do any of those who entertain and incorporate in their own thinking the challenges of the 1980s critiques claim resolutions of them toward new paradigms or models of practice. There are many demonstrations why realist ethnography cannot return in its previous forms, but few of the authors worry about alternatives other than those that each is developing within the confines of his or her own text and experience of fieldwork.

What the 1980s critiques of anthropology accomplished was to evacuate the space defined by ethnographic conventions so as to license experimentation within limits. The critiques were shaped by questioning the conventions that defined the authority of ethnographic knowledge in terms of its discursive expression and rhetoric, by encouraging hyper-reflexive modes of thought and writing, and by establishing representations as social facts and their production as the primary focus of the eth-

nographic analysis of social process. Now there is ambiguity, amply expressed in these volumes, about whether any alternative (or counter) authority can be asserted, and if so, on what terms (e.g., Can anthropology now have confidence in the validity of its own questions?). The stronger papers move toward new bases for the authority of ethnographic inquiry but through, rather than around, the difficult issues of producing authoritative knowledge on any human subject marked as "Other."

Reflexive writing is the pervasive characteristic of the pieces in both volumes. Reflexivity as generative of new research strategies rather than as a means of self-discovery is most strongly developed in the pieces of *Auto/Ethnography. After Writing Culture* tends to develop representation as much as an object of study as a reflexive means of study. It usefully extends the question of representation beyond issues of ethnographic writing within textual genres to the social milieus in which ethnographic research is embedded, from fieldwork through to the reception of anthropological writing by various constituencies. The fact that anthropology no longer discovers subjects who are not always already written about calls for a rethinking of the identity and function of anthropological analysis in the complex, preexisting fields of representation that these volumes only begin to broach.

Those who thought the critiques of the 1980s were too negative and a serious diversion of limited duration from the positive pursuit of knowledge will be disappointed to learn that the critical turn continues to be explored even more deeply by the most talented contemporary anthropologists represented in these volumes. Still there is much to be worked out in this continuing exploration of the problems of authority, reflexivity, and representation within ethnographic practice. What is often pointed at, but not explicitly stated, through these pieces are arguments that would shift the very conditions and conventions that define the imaginary of ethnographic research. The papers that go farthest in this direction are those that define "mature" work, those that are transformations of, and more complex layerings upon, earlier projects of ethnography in the same places. Kay Warren discusses Pan-Mayan intellectuals, who, writing as activists and scholars, have to some extent displaced traditional ethnography, its functions, and rationales. An established ethnographer of Mayans under military repression, Warren provides a searching, complex consideration of the options for outside anthropologists so displaced who now can only continue to relate to Mayans through their own intellectual movements. Michael Herzfeld employs the current modes of reflexive critical writing to provide a stunning integration of his earlier distinctive ethnographies in Greece among the milieus of bureaucrats, sheep thieves, and a novelist. He shows the value of these modes for mature scholarly reflection on a career of ethnographic soundings into the puzzles of national identity.

In a sense, then, the 1980s critiques of anthropology have been most importantly developed by those who have kept working through the changing personal and societal histories that have affected the places of

their earlier ethnography, done before the critiques settled into disciplinary debates. A challenge for the future is how the wisdom of these mature projects can be incorporated in the training models by which anthropologists in the making produce their first works. This will require going further than these papers have toward evolved and modified practices of ethnography that might define new norms of practice for the disciplinary community of anthropologists.

Finally, while the editors of these collections do an admirable job of defining issues in their introductions, the pieces themselves often fall short of the theoretical or conceptual development that their discussions need. This lack derives from both the virtue and limitation of the habit of anthropologists to write strictly within the terms of their ethnographic cases. The virtue is that these pieces are always rich in setting, always about something specific and substantive. Their limitation is connected to this same virtue—writing within the frame of a particular ethnographic experience relieves one of the inclination or responsibility to develop implication, to argue beyond the confines of ethnographic particularism. The result is that the struggles of a discipline going on within the predicaments of individual ethnographic projects might be missed or even misunderstood as a discipline stalled on piecemeal critical introspection when in fact there is, perceptible in these volumes, an adaptation underway toward new practices keyed to the changing conditions of research and the nature of anthropologists' relationships to their subjects of study. In this light, most of these papers are about the search for new formulations of the anthropologist's voice and position as a shared, disciplinary matter about which a vigorous, and mostly constructive politics continues to unfold.

Redeeming Laughter: The Comic Dimension of Human Experience. By Peter L. Berger. New York: Walter de Gruyter, 1997. Pp. xvii+215. \$23.95.

Mahadev Apte
Duke University

This book is another addition to the burgeoning literature resulting from a resurgence in humor research that has taken place during the last 20 years or so as manifested by the publication of many books, several national and international humor conferences beginning in 1976, and the establishment of the International Society for Humor Studies and of the international humor research journal *HUMOR*, which is now in its tenth year of publication.

In addition to "Prefatory Remarks, Self-Serving Explanations and Unsolicited Compliments" and a prologue, there are 14 chapters divided into three parts. A poem titled "The Donkey" by C. K. Chesterton serves as the epilogue. Berger prefers to use the term "the comic" instead of "hu-

mor" throughout this work since he claims that the comic is "the objective correlate of humor," which he considers to be "the subjective capacity." He also claims that "the comic is experienced as *incongruence*" (p. x).

Chapters 1–6 in part 1 focus on the anatomy of the comic. Chapter 1 identifies and elaborates on the concept of the comic and its similarities to and differences from dreaming, the experiences of the esthetic and of sexuality, and play. The contributions of ancient and modern philosophers from Plato to Plessner are briefly discussed in chapter 2, while chapter 3 is devoted to a very brief discussion of the Chinese and Japanese philosophy of the comic. Chapter 4 discusses the physiology and psychology of the comic, and chapter 5 critically examines comedy, carnival, circus clowns, court jesters and fools, and medieval folly, which are, in Berger's view, the "social constructions of the comic" (p. 65). Chapter 6 reflects on the nature and *raison-d'être* of Jewish humor.

Chapters 7–10 in part 2 discuss benign humor, tragicomedy, wit, and satire, respectively, while chapter 11 defines and elaborates the concept of the absurd and its connections to the comic. The notion of folly in its modern version is also briefly reexamined here. The writings of P. G. Wodehouse, Sholem Aleichem, H. L. Mencken, Oscar Wilde, Will Rogers, Karl Kraus, Jarry, Franz Lehar, and Eugène Ionesco are used throughout this part for illustrative purposes.

Chapters 12–14 in part 3 explore the religious implications of the comic, which Berger considers to be "a theology of the comic" (p. 185). He explores the notions of the holy fools and holy folly within the context of Old and New Testament with special emphasis on several religious figures, including Jesus. For Berger the experience of the comic is a signal of transcendence, and he believes that the comic is one manifestation of a sacramental universe that "contains visible signs of invisible grace" (p. 214). The comic, if perceived in faith, "becomes a great consolation and a witness to the redemption that is yet to come" (p. 215), hence the "redeeming" value of laughter.

Berger wrote this book because "the comic is a subject that cries out for unprofessional treatment" (p. ix), and he would prefer to consider his work as "a prolonged reflection about the nature of the comic as a central human experience" (p. x). Like other books on the subject, it highlights the multi- and interdisciplinary nature of humor research. However, its problem is that Berger tries simultaneously to be selective, reflective, and also comprehensive in his treatment of the comic. Thus while Berger cites extensively his favorite scholars such as Zijderweld and Schutz, highlights his favorite topics such as Jewish humor, and fools and folly—the latter being significant for his discussion of religion and humor—and ruminates about the nature of the comic, he also attempts to "explain" the comic from perspectives ranging from the natural sciences to humanities in a rather cursory manner.

While Berger's extensive foray into a "Theology of the Comic" (pp. 185–215) is quite illuminating, it is not clear who his book's audience is. If it is meant for humorologists, there is little new in it. If it is for

other scholars who may wish to know what is going on in the field of humor research, despite its publication in 1997, the book is not cognizant of the substantial theoretical and methodological developments that have taken place in many disciplines and reflected in several major publications. Whether this lack of awareness is deliberate because Berger wants his book to be an "unprofessional treatment" or it is unintentional, one cannot say. But since Berger cites many works and references despite his unprofessional stance, one would expect that he would at least show an awareness of the tremendous scholarly output during the last two decades. If the book is for educated lay persons, they may find it quite interesting because of Berger's engaging style, his easy-to-follow analytical explanations, his humorous comments and asides in texts and footnotes, and the many jokes interspersed throughout the first two parts. There is no index despite a claim to the contrary on page iv.

Misery and Company: Sympathy in Everyday Life. By Candace Clark. Chicago: University of Chicago Press. 1997. Pp. xii+316. \$29.95.

Jack Katz
University of California at Los Angeles

Candace Clark's *Misery and Company* is a very useful intellectual project. Frequently exploiting George Eliot's fiction for insights, Clark's book suggests a rare, almost 19th-century curiosity about a familiar institution in everyday social life. Despite the academic presentation of the evidence, the focus remains rooted in an inspiring wondering about the significance of sympathy.

And there is plenty to wonder about. People not only confer sympathy ubiquitously in everyday social life, they are highly sensitive to several collateral matters. Sympathy givers assess the demands that others make for expressions of sympathy, honoring those who make few demands and critically viewing those who make too many. As Bourdieu has made clear, recipients often resist the debilitating implications that come with unsolicited gifts of sympathy. Men and women give and receive sympathy differently, and the differences are significant enough to indicate that something major about gender identity is involved. In the form of commercialized greeting cards, "sympathy entrepreneurs" have exploited interactional needs to create very big business.

Clark uses several analytical perspectives to sort out and reveal subtle dimensions of her data. She considers sympathy as ritual and moral drama, runs through a functional analysis of how sympathy giving contributes to social order ("social glue" theory), and works with an economic metaphor that casts recipients as saving and using up sympathy credits, a metaphor that she relates to Gouldner's "norm of reciprocity." Considering the phenomenon of sympathy in its own right, there appears to be a distinctive logic to expressions of sympathy. They imply that something

negative in the recipient's experience should be attributed to bad luck, that the recipient and the donor of sympathy share some important aspect of personal identity and are related in some kind of ongoing social bond, and perhaps that the recipient merits a relaxation of role obligations.

Most intriguing is her treatment of how expressions of sympathy provide a sense of "place," a kind of footing for action that is collaboratively maintained, and collaboratively repaired when trouble is sensed or announced. "Sense of place arises in a particular interactional context. . . . It is the momentary consciousness of 'who I am and how I can act at this moment in this encounter'—part of the situated self" (p. 230). Requests for and grants of sympathy are directly relevant to the operative sense of "place." Clark's pioneering step here suggests that sympathy, and the study of emotions more generally, is essential to the analysis of all forms of social interaction.

Clark's use of evidence puts the book at a critical juncture in the sociological study of emotions. She seeks to mesh survey, ethnographic, and even psychodynamic forms of evidence, and she uses negative cases particularly well (e.g., the lack of sympathy that some felt for Anita Hill) to bring out hidden preconditions of the emotion. In a useful appendix, Clark describes her struggles to match methods with the special demands of her topic. Subsequent work on sympathy, and indeed on emotions in general, is likely to divide on the issue of whether the distinctive quality of emotional experience requires special methods.

With Tom Scheff and Carolyn Ellis, Candace Clark believes that because emotional experience is both personally felt and socially presented in routinized patterns of expression, we need special methods to get underneath observable, socially presented conduct. After all, feelings and their representation to others are not necessarily the same thing. How are we to avoid the errors of, on the one hand, assuming that what we see is what the other feels, or, on the other, imputing feelings that go beyond what is observable in our data? In ways that are reminiscent of psychological forms of inquiry, Clark does "distanced readings" of novels, conducts intensive interviews that seek to get respondents to describe their feelings, seeks "Piagetian" responses to experimentally varied vignettes, and has subjects do "free writing."

The premise of calling for special methods to study emotions is the familiar assumption that emotions entail something "inner" that is not available in the "outer" layers of observable conduct. But that assumption is not inevitable. An alternative view is that, while there is a variable relationship between the corporeal foundations of one's behavior and what one represents about the self for others to witness, both are visible to the researcher. If the person in social interaction must keep the felt, bodily grounds of conduct invisible to self-reflection and masked from others' immediate observations, the practices of creating this ongoing invisibility may still be documented by analysts. This is because analysts need not act and respond in the real time and original place in which conduct occurs. We have luxuries that our subjects lack. From the side-

lines, we can write detailed field notes on comportment and expression. After interviews, we can comb subjects' detailed narratives of prior experiences without immediately going with or against the moral of their story. And we can repeatedly review videotapes of naturally occurring, socially situated behavior. In these ways we may reach "deep" or qualitatively distinctive aspects of emotional experience, not by some metaphysical trick by which we probe "within" subjects' minds or experiences, but by observing the practices that people carefully keep in the background as they produce expressions of self to take effect on the foreground of social interaction.

Risk and Misfortune: A Social Construction of Accidents. By Judith Green. London: UCL Press, 1997. Pp. x+222. \$26.95.

Joseph Gusfield
University of California, San Diego

In one of his essays on punishment and responsibility, the philosopher of law H. L. A. Hart remarked that the Ming vase was just as lost to its owner whether as a result of theft, intentional destruction, carelessness, or accident. Yet the response of society was very different toward each alternative. What is construed as an accident? What consequences follow upon differing perceptions and conceptions of an event as accident? How do these vary in specific historical periods? These vital questions are considered in this important and excellent book.

In *Risk and Misfortune* Judith Green, lecturer in sociology at the London School of Hygiene and Tropical Medicine, examines how it is that societies develop and utilize answers to explain the occurrence of misfortunes. The concept of "accident" is not a universal one, found at all times and in all societies. Its appearance and its usage is not a natural product but owes its existence and its meaning to social construction.

Many decades ago Job asked the question still so salient in contemporary eras: Why me, oh God? Green finds that the concept of "accident," as a random or inexplicable event, emerged in the West toward the end of the 17th century as a facet of modernity. Prior to that time such misfortunes were seen as outcomes of Fate, Providence, witchcraft, God's will, or similar explanations. With modernity the concept of "accident" implies a residual category of events that cannot be explained; that "just happen."

As the Freudians are wont to say, "it is no accident" that some events came to be seen as "accidental." The modern concept of accidents as unexplainable events, at the margin of determined actions, depended on the emergence of rationality and the idea of probability. Since no rational explanation of the event is able to be made and nonrational ones are unacceptable, the specific and particular event is "accidental." "The modern belief in a rational universe would hardly have been sustainable with-

out a belief that some misfortunes just happen. . . . The accidental in social life appears from the events deemed worthy of no further inquiry; the remnants of our explanatory system" (p. 67).

But cultural life is not as static as it sometimes seems. The concept of "accident" as a residual category has given way in the late 20th century (in what Green calls "high modernity") to the idea of risk. Risk relates misfortune to general categories, eschewing the explanation of the particular event but explaining and predicting the general. Thus cigarette smokers are seen as more exposed to greater risk of early death from lung cancer than are nonsmokers. But which smokers and which nonsmokers will be victims is still unexplainable by the generalizations. Job's question remains. As every craps shooter knows, the general probability cannot predict the particular throw. The concept of risk, however, has given rise to the idea of risk prevention; that we can by our efforts minimize and avoid risks which pertain to the generalized categories of membership. The smoker can quit. Risk is our modern theodicy.

Despite the amoral character of the concept of "accident," the moral character of misfortune is not to be lost. In a chapter based on her work with focus groups, children and adult, Green provides many examples of interpretations of individual, situated misfortunes deemed accidents but replete with attributions of fault. The concept of accident persists as a substantive category, as in "auto accidents," but the moral attributions are nevertheless clear. Whatever the contribution of modernity, the epistemology of accident and the statistical probabilities of risk seem belied in the talk of her respondents.

It is the mark of a good piece of research that it raises questions and opens lines of thought that were nonexistent or dormant. This *Risk and Misfortune* does. In what is ignored we can find new avenues of worth down which to walk. She touches only peripherally on the very important question of the role of the idea of insurance in modern society—a question to which sociologists, except for Carol Heimer, have paid little attention. The legal issues over strict liability and fault, set deeply in the concept of risk and risk allocation, are untouched. Green's work is based on a vast bibliography but she has made scant use of some significant sociological predecessors, such as Perrow's 1984 study of normal accidents, and no use, alas and perhaps self-servingly, of my 1981 study of drinking and driving.

In the usual disclaimer, despite some flaws, this is an exciting book with many implications for future thought and study about contemporary society, its continuity and discontinuity with the modern, the premodern and the postmodern. Yet despite all the change, Job's question remains: Why me, oh God, why me?

Making Crime Pay: Law and Order in Contemporary American Politics.
By Katherine Beckett. New York: Oxford University Press, 1997.
Pp. vi+158. \$27.50.

David Jacobs
Ohio State University

There is much to like about this brief book, but it has a few faults. Katherine Beckett focuses on the politics of crime control and shows how her political approach can explain some important issues in criminology and political sociology. The recent dramatic growth in punitive reactions to crime has left huge numbers of young black males in prison. Public resources devoted to corrections and law enforcement have increased to the point where they are beginning to crowd out expenditures on valuable programs such as education in many jurisdictions.

Beckett begins with an intriguing time-series analysis of the factors that lead to shifts in the salience of the crime issue, but her main interest is in describing recent changes in the public discourse about crime. Her time-series findings are fascinating, but she could have provided more information about methods. After holding the crime rates and media coverage constant, Beckett finds that pronouncements by national political officials help determine the degree to which the public is concerned about crime as revealed in Gallup polls. If that finding is correct, it is significant. One reasonable inference is that much of the public's anxiety about law and order is manipulated by politicians who increase their support by making this issue more conspicuous. If these statistical results can be trusted, it would be reasonable to suspect that a good part of the rapid growth in incarcerations since the early 1970s and the attendant acceleration in expenditures can be attributed to partisan tactics by conservative politicians.

The devil is in the omitted details. Significance tests that determine the theoretical implications of time-series findings generally rest on corrections for autocorrelation. Researchers who do not use equal time periods will find that it is difficult to eliminate this threat to the external validity of their results, but Beckett does not tell us how she handled these problems. In time-series designs with lagged explanatory variables, as long as autocorrelation is eliminated and the other standard requirements for regression hold, reverse relationships from the dependent back to the independent variables are unlikely to bias the estimates. Yet Beckett thinks her coefficients may be inflated by this difficulty, so I am unsure about what she did. After she took so much trouble to code the contents of newspapers and political speeches, it is too bad that readers are left uncertain about the accuracy of these stimulating regression results.

The rest of the book concentrates on the literature about the relationship between politics and public perceptions of crime. Chapters 2 and 3 present a political history of criminal justice policy that primarily focuses on what happened at the national level since the 1960s. In chapter 4

Beckett uses her content analysis of media coverage to develop a useful classification of media depictions of crime and the (often implicit) remedies that follow from the writer's presuppositions.

In chapter 5 she presents a review of the literature on the associations between racial attitudes and opinions about how crime should be controlled. In support of her thesis, the available studies suggest that people who are likely to endorse punitive responses to street crime also are likely to hold unfavorable opinions about African-Americans. Yet groups who are especially anxious about crime are not as likely as others to support harsh measures. The aged and nonwhites are most concerned about crime, but white males more often endorse punitive remedies.

Another strength is the diverse literature Beckett uses to address these issues. She finds quotes from officials in the Nixon administration stating that they deliberately went after the racist vote with rhetoric about street crime. She uses material from the Edsall's compelling account of the recent politics of race (*Chain Reaction: The Impact of Race, Rights, and Taxes on American Politics* [Norton]) that more sociologists should read. Political sociologists have neglected the recent appearance of conservative populism and its emphasis on race and street crime. That neglect has left us unable to explain the outcomes of most of the presidential elections since the end of the 1960s. Beckett's work suggests that sociologists who study politics might find it useful to look more closely at the political uses of the crime issue.

This book would be a profitable assignment for advanced undergraduates or graduate students. In contrast to much of the literature, it spells out some links between social theory and criminology. It also shows how useful an informed combination of political sociology and criminology can be. I therefore recommend it to researchers as well as to students.

Penitentiaries, Reformatories, and Chain Gangs: Social Theory and the History of Punishment in Nineteenth-Century America. By Mark Colvin. New York: St. Martin's Press, 1997. Pp. x+294. \$45.00.

Thomas F. Courtless
George Washington University

The author sets out to determine the extent to which the transformation to a market-oriented society shaped systems of punishment in 19th-century America (p. 5). He applies four "rival" theoretical perspectives to his task: Durkheim, Marx, Foucault, and Elias. This work clearly is relevant in the context of the current debate surrounding the roles and rationales of punishment and, especially, the role of incarceration.

The work consists of three case studies: the rise of the penitentiary in the Northeast, the transformation of gender roles and the punishment of women offenders in the North, and the transformation of criminal punishment in the South. Each case study contains a concise history of the

social and economic context in which penitentiary and other institutional forms of punishment developed. Following each case study is a chapter examining critically the four perspectives for their explanatory value.

Missing from the case studies are references to Newgate of Connecticut, America's first prison, and Dorothea Dix's survey of prisons first published in 1845 (see *Remarks on Prisons and Prison Discipline* [Patterson Smith, 1967]). Also, a discussion of the Irish system that directly influenced the American reformatory is not provided.

The author points out a fundamental contradiction confronting corrections: control versus rehabilitation. As an example of this he says that "the goal of the Auburn system was to break the spirit of the prisoner and produce a mental state of complete submission. . . . Whether the prisoner could be reformed after being brought to this state was a matter of dispute" (p. 91).

Throughout the work, the author highlights periods of pessimism in the attitude of many elites toward incarceration. "This pessimism is the penitentiary's most devastating cultural product; it . . . reinforces an ideology of futility that undermines our attempts to create a society based on economic and social justice" (p. 127).

In the case study dealing with the punishment of women, the author presents a thorough discussion of the importance of the ideal of "true womanhood" and the "cult of domesticity" as these shaped women's reformatories and the confinement regimes in them. Reforms that were introduced into women's prisons and helped break them away from the model set by those for men stressed the inculcation of "desirable American virtues for young women" (p. 149). One targeted population was Irish immigrant girls who seemed to violate the ideals of true womanhood because of the support Irish culture gave women who played roles outside the home, roles considered gender inappropriate in America. At the same time, indeterminate sentencing had the consequence that women served longer terms than men for the same or similar offenses.

In the third case study, the author argues that there developed a "unique" southern system of punishment that evolved out of a history of private justice. This was especially true with regard to black offenders. Plantation-based justice preceded and largely influenced public justice following the Civil War. The development of convict-leasing systems and chain gang sentencing, with their relatively minimal reliance on state prisons, left prisoner control largely in the hands of private contractors. Black offenders were frequently leased out to private "employers" who paid their fines.

In examining the theoretical perspectives in relation to southern punishment systems, the author tests Elias's focus on "sensibilities" in the North and South. In the former, it was "civilized sensibilities" that influenced the penchant for state prisons, while in the latter it was "honor" that was at stake. Thus, in the South, the sense of personal honor supported private justice and opposed centralized prisons. Neither Elias's nor Foucault's perspectives are of much help: "What is striking about

the case study of southern punishment is the absence of factors postulated by Foucault and Elias. Neither sophisticated, rational systems of punishment nor civilized sensibilities play significant roles in the transformation of punishment in the South" (p. 264).

Examining the "differences in the sensibilities of the elites" in the North and South, the author found that "the moral force of white supremacy combined with the profit motive to shape the southern penal system" (p. 265).

At the end of the work, the author discusses the legacy left by 19th-century punishment systems. These include boot camps, supermaximum security prisons, and chain gangs.

The author concludes that Durkheimian and Marxian perspectives appear to offer more than Foucault's or Elias's in understanding contemporary punishment systems and practices. Government at all levels is failing to confront the underlying causes of criminality, as historically has been the case, focusing instead on the "drama of crime and punishment" (p. 273).

It is difficult to know where this work fits in an academic setting. The author wrote it to provide material for his courses in the sociology of punishment. Certainly it belongs there, although there may not be many such courses out there.

Casualties of Community Disorder: Women's Careers in Violent Crime. By Deborah R. Baskin and Ira B. Sommers. Boulder, Colo.: Westview Press, 1998. Pp. xiv+192. \$59.00 (cloth); \$22.00 (paper).

Pat Carlen
University of Bath

Based on in-depth interviews carried out between 1992 and 1994 with 170 women who had committed violent felony crimes in New York City, *Casualties of Community Disorders* offers a theory of women's violent crime that, its authors claim, answers two distinct questions: "How are women's criminal careers . . . circumscribed by gender? And what accounts for the increasing pull of women into the world of violent crime?" (p. 7).

The answer they give to the second question is that the violent crime careers of contemporary women in New York City are shaped by many of the same factors as those of their male counterparts: primarily, the culture of particular communities, and specifically the culture of drug-dealing communities where enforcing respect (through violence if necessary) is as important as the dealing done to fund the drug habit; and, additionally, the influence of peer groups, the quest for excitement, illicit opportunity structures and other neighborhood effects. For criminologists, there is little that is very original here. Nonetheless, in proffering such an explanation, Baskin and Sommers do manage to "bring women

back in" to mainstream criminological explanation. What they fail to do is answer their first question and explain how women's criminal careers (and especially women's *violent* criminal careers) are circumscribed by gender.

In fact, gender is hardly mentioned in the book, and nowhere is there any discussion of, first, how the effects of gender on crime might be theorized; and then, secondly, how that theorization might be operationalized in an empirical research project such as the one the authors undertook. Indeed, at the end of their book, they conclude that there is a "new dynamics of crime in which gender is a far less salient factor" (p. 150). Maybe so—in *some types of crime*—but, given the question about gender that Baskin and Sommers themselves pose, that negative finding in relation to gender does not excuse the lack of any theoretical discussion of what the possible relationships between gender, women, and crime might be. Nor does it excuse the writers' apparent equation of "gender" with "women."

There are other disappointments, the main ones relating to the underrepresentation and undertheorization of the considerable amount of interview material amassed. Only 23 pages of the book are specifically about violence, and then the voices of the women interviewed are heard sparingly; again, readers are not given much indication of the richness of the data that the authors collected. When quotations from the women are used, they tend to be mainly descriptive, rather than reflective, and the authors do not present them in order to build any new theory of their own about women, violent crime, and gender. Rather, they tend to use them to illustrate already-known theories explaining men's crimes. They might have been able to make more and better use of their extensive data if they had spent less time tilting at the theoretical windmills supposedly erected by all previous writers in the field.

In addition to their two main questions, the authors, at the beginning of their book, pose all the nowadays conventional questions about the relationships between class, age, ethnicity, and gender. Unfortunately, and despite their claims to the contrary (p. 24), they do not, thereafter address these actual/possible relationships either empirically or conceptually. Instead, they merely describe each factor's separate impact on the women interviewed.

Yet, despite its theoretical omissions, this is a very readable and vivid account of life in drug-dealing communities. The book is well structured around the conceptualization of different stages in the women's criminal careers, and the authors conclude with a discussion of the policy implications, which is insightful and important.

Sociology and criminology students could be expected to find this study a riveting and pleasantly digestible introduction both to basic criminological theory and to research methods. Academic criminologists will find the authors' approach refreshingly lacking in the ideological baggage that has at times surrounded women-and-crime issues; and social scientists

and policy makers should find the study itself a very welcome addition to the literature on urban drug-dealing communities.

Overall, then, there are enough good things in *Casualties of Community Disorders* to make this reader, at any rate, hope that the authors will hereafter use more of their data in a second volume and then answer some of the very important questions which they raised, but failed to answer, in the first.

A Cultural Field in the Making: Gastronomy in 19th-Century France¹

Priscilla Parkhurst Ferguson
Columbia University

The gastronomic field in 19th-century France is taken as a model for the analysis of cultural fields as characteristically modern phenomena. The antecedents of the field are located in a new economic, institutional, and ideological context. But its foundations are laid by a spectrum of gastronomic writings (journalism, cookbooks, proto-sociological essays, political philosophy, and literary works) that proposed an expansive, nationalizing culinary discourse. It is this discourse that secured the autonomy of the field, determined its operative features, and was largely responsible for the distinctive position of this cultural field.

CULTURAL FIELDS

Although it has been applied to many enterprises, the concept of "field" has proved especially fruitful for the analysis of intellectual and cultural activities. Elaborated in its specifically sociological usage by Pierre Bourdieu ([1966] 1969), "field" designates the state of a cultural enterprise when the relevant productive and consumption activities achieve a certain (always relative) degree of independence from direct external constraints (i.e., those of state and church for the arts in premodern Europe). As a "particular social universe endowed with particular institutions and obeying specific laws," a field translates external economic or political phenomena into its own terms for its own use or, rather, for the use of its occupants (Bourdieu 1993, pp. 163–75). To the extent that the norms governing conduct, the values inducing behavior, and the rewards determining produc-

¹ It is with real pleasure, both intellectual and personal, that I express my appreciation for the insights, the sociological sense, and the larger understandings offered by Robert A. Ferguson, Jeffrey Olick, and Harrison White. I am also grateful to the *AJS* reviewers for the strategic reformulations that their criticisms prompted. Special thanks are due Sharon Zukin for urging and inspiring me to join our forces in culinary research. Please direct all correspondence to Priscilla Parkhurst Ferguson, Columbia University, Department of Sociology, MC 2555, 1180 Amsterdam Avenue, New York, New York 10027. E-mail: ppf1@columbia.edu

tion operate according to field-specific standards, a field is self-regulating, self-validating, and self-perpetuating.

Thinking in terms of cultural fields modifies our understanding of cultural enterprises. Against the functional divisions that tend to be drawn for such activity, a field constructs a social universe in which all participants are at once producers and consumers caught in a complex web of social, political, economic, and cultural relations that they themselves have in part woven and continue to weave. Against unilinear, univocal approaches that focus on discrete structures, historical incident, or individual producers and products, the complex, dynamic configuration of social and cultural relations proposed by a cultural field offers a model that can do justice to the many and diverse modes of cultural participation on the part of a broad range of individuals, institutions, and ideas. The foundations of the cultural field are laid by neither the singular cultural product nor the producer but by a spectrum of products and practices that displays the workings of the field as a whole.

Cultural fields have an advantage over encompassing sectors like politics or the economy in that they focus our attention on tangible products and identifiable pursuits. A sustained concentration on cultural fields—their internal disposition as well as their external relations—stocks the sociological arsenal with the kind of controlled studies that integrate empirical, historical evidence into a conceptual framework. The more circumscribed the field, the more solid the ground for sociological scrutiny. It is not surprising, then, to find that the most successful studies work with the specifics of a given sphere of cultural production: the “literary field” proposes a delimited space for investigation; a vast, necessarily imprecise construct like the “field of power” invites speculation.¹ The more limited focus facilitates situating the field as a historical entity as well as a sociological concept. The analysis below also demonstrates that this particularity of focus also furnishes useful analytical distinctions between the related but distinct notions of “field,” “culture,” and “world”—all of which have been invoked in contemporary sociological discussions, particularly, although not exclusively, for the arts.

A sharper use of the concept of “cultural field” and the power to focus inquiry go far to account for the specifically sociological interest of gas-

¹ The difference is evident in Bourdieu's own work. See, in addition to the suggestive discussion of the journalistic field in *On Television* ([1996] 1998), the extensive empirical examination of *Homo academicus* ([1984] 1988) and the grounded interpretation in *The Rules of Art* ([1992] 1996). Typically, Bourdieu's analyses navigate between more and less closely defined fields (*The Field of Cultural Production* [1993, chap. 5]). Alain Viala's examination of the 17th-century literary field in France (1984) works so well precisely because it scrupulously centers on readily observable cultural products and practices.

tronomy in 19th-century France. As a relatively delimited cultural enterprise, the pursuit of culinary excellence that we call gastronomy enables us to address a number of problems that plague discussions of cultural fields. Most notably, it speaks to the sticky issue of antecedents. For, however good an idea we may have about how certain fields operate, we know rather less about how they got to be fields.³ It is true that any search for “causes” or even “origins” is doomed to fail. Yet, the question must be put: At what point do structures and sensibilities, institutions and ideologies, practices and practitioners cohere to “make” the configuration that we designate a cultural field? To this question, gastronomy proposes some answers. For, although the culinary arts in the West can be traced to the Greeks and especially the Romans, gastronomy as a modern social phenomenon was instituted in early 19th-century France. It was then, I argue, that the culinary arts moved into public space and acquired a public consciousness that justifies identification as a “gastronomic field.”

Gastronomy turns out to be a happy choice. On the one hand, it speaks to the broad controversy over the meanings of modernity, and, on the other, it addresses an issue that is unavoidable in almost any discussion of 19th-century French society, namely, the real or supposed effects of the Revolution of 1789. In what sense can a given cultural venture be considered “modern”? If debates over modernity and modernization, as well as assertions about cultural fields, tend to assume that the visibly changed and changing society of the 19th century favored the separation of cultural enterprises into relatively distinct and autonomous domains, it is not at all clear how this transformation occurred. To evoke literature for a moment, in what ways does the Republic of Letters—a term that recurs regularly in 18th-century French intellectual life as a designation for networks of writers and thinkers—differ from the literary field of the 19th? For the culinary arts, how is 19th-century gastronomy “modern,” or distinct from elite culinary practices in the 17th and 18th centuries?

If the paradox of eating, as Simmel ([1910] 1994) pointed out in a quirky but suggestive piece, is that this physiological activity shared by every human being should give rise to such extraordinary social differentiation,

³ Bourdieu's own empirical analyses tend to map the field as constituted. They mostly draw the larger societal and intellectual consequences in order to identify the mechanisms and the logic by which the field reproduces itself. *Homo academicus* (1988) begins with the postwar university field and gives relatively little consideration to the conditions out of which the field emerged. Although the subtitle of *The Rules of Art—Genesis and Structure of the Literary Field*—indicates the goal of tracking the emergence of the literary field in mid-19th-century France, the title betrays the theoretical scope. Moreover, the emphasis falls on the structure and logic of the literary field and its evolution in the last half of the century rather than on the conditions out of which the field emerged in the first half.

it is clear that modern gastronomy enriched the social order by exacerbating those distinctions. Gastronomy constructed its modernity through an expansive culinary discourse and, more specifically, through texts. Gastronomic texts were key agents in the socialization of individual desire and the redefinition of appetite in collective terms. The "second-order" culinary consumption of textual appreciation was as crucial for the construction of the gastronomic field as it was (and is) for its operation. Such writings extended the gastronomic public or "taste community" well beyond immediate producers and consumers. Diners, thus converted into readers, became full-fledged participants in the gastronomic field. The public sustained the gastronomic field, and the field determined the public. As with the performing arts, writing about food presupposes a different order of consumption inasmuch as the cultural product in question is at one remove from the base product—the work performed, seen, or heard and, in this instance, the food prepared and consumed. These culinary texts of indirection were indispensable for the gastronomic field because they stabilized the ephemeral culinary product within a network of nonculinary discourse and because they redefined the culinary as broadly cultural. Texts, both instrumental and intellectual, are therefore critical in making food what Mauss (1967, p. 1) identified as a "total social phenomenon"—an activity so pervasive in society that, directly or indirectly, it points to and derives from every kind of social institution (religious, legal, and moral) and every type of social phenomenon (political, economic, and aesthetic). That food so penetrates the social fabric is the work of many factors. But pride of place surely goes to these texts and writings. To turn singular food events into a veritable cultural configuration, to transform a physiological need into an intellectual phenomenon, dictates powerful vehicles of formalization and diffusion. The gastronomic writings that proliferated over the 19th century supplied the mechanisms that brought the culinary arts into modern times.

In France, reflections on modernity further necessitate coming to terms with the Revolution of 1789. What responsibility for the institution of a recognizably modern social and cultural order can be ascribed to the many and varied phenomena associated with the Revolution and its immediate consequences?—the abolition of the monarchy, the elimination of traditional economic constraints on commerce, the foreign wars and domestic political turmoil, to list only the most obvious elements.⁴ Given that the

⁴ At the very least, 19th-century France was the site of considerable turbulence. From 1789 to 1871, there were three monarchies, three republics, and two empires; three revolutions (1789, 1830, 1848), one coup d'état (1851), and one insurrection (the socialist Commune of 1871). Napoleon I's defeat at Waterloo in 1815 ended almost a quarter century of war and put France under occupation; Napoleon III's devastating defeat

theoretical model of the cultural field in no way demands that all fields have the same degree of coherence or follow the same logic, the connections of cultural fields to both modernity and the Revolution argue for comparative analysis, across fields as across societies. The sociological issue then becomes the identification of those factors that distinguish gastronomy in France as a historical phenomenon and as a cultural practice. To what degree is this field anchored in, and therefore definable in terms of, distinctive cultural traditions and particular historical circumstances?

The gastronomic field took shape in two major phases: emergence over the first half of the 19th century, consolidation thereafter. The resulting cultural formation carried "French cuisine" well beyond a circumscribed repertoire of culinary products to comprehend the practices and products, values and behavior, rules and norms, institutions and ideas that are attendant upon the preparation and consumption of food in this particular social setting. The gastronomic field turned a culinary product into a cultural one. This cuisine became "French" as it had not been in the 17th and 18th centuries when the culinary arts were associated with the court and the aristocracy, not the nation. Culinary institutions and texts in the 19th century effectively transformed the patently class-based culinary product and practices of the *ancien régime* into a prime touchstone of national identity.⁵ The consequent identificatory power of cuisine as a fundamental attribute of "Frenchness" and the high rank of the gastronomic field in the hierarchy of cultural fields in France are a function of the strength, the extent, and the multiple and prestigious associations of the gastronomic field, notably its many and varied affiliations with the literary field.

Five structural factors signal the transformation of gastronomy into the gastronomic field. First, new social and cultural conditions stimulated production, sustained broad social participation, and encouraged a general cultural enthusiasm for the product in question. Second, specific sites came to be dedicated to cultural production and consumption. Third, the institution of standards and models of authority ensured an acute critical consciousness that focused and checked yet also legitimated the expressions of cultural excitement. Fourth, subfields generated by continued expansion of the field assured the simultaneous concord and conflict of the parties involved, the consonance and dissonance of new positions and

by the Prussians in 1870 after a mere six weeks of military engagement led to a second occupation by enemy troops as well as significant loss of territory (Alsace-Lorraine).

⁵ Arguably, only the culinary class changed, with the bourgeoisie replacing the aristocracy. But the (self-)identification of the bourgeoisie with the nation in this as in other cultural and political concerns effaced the class connotations of cuisine and implanted patriotic ones through the kinds of discourse discussed below. See also Ory (1992).

alliances. The resulting interlocking networks of individuals and institutions forged links with adjacent fields, and it is these linkages that were largely responsible for the social prestige of gastronomy.

As the formulation of these field qualities suggests, I conceive of gastronomy in 19th-century France as something of a template for the analysis of cultural fields more generally. Its lessons reach well beyond the kitchen and the dining table. Certainly, to appreciate a cultural field in the making is to grasp the concept and its use in a more rigorous way, but to do so raises an intriguing problem. To actually see a cultural field in the making requires a delicate conjunction of historical and sociological recognition. The emerging historical phenomenon and social structure must be apprehended, in a word, simultaneously.

The discussion below begins by locating the foundations of the gastronomic field in France in the complexity and the confluence of institutions, traditions, attitudes, events, and ideas. A second section then deals with the articulation of these phenomena in what I have termed culinary discourse. Without such a discourse, I argue, there can be no cultural field. The sociologist, then, must pay careful attention to the specific terms of the discourse. A third section and conclusion assess the validity of the gastronomic field as an analytic category by considering it against other related cultural fields. We shall find here that the vitality of this particular cultural field depends absolutely on its relations to other fields.

FOUNDATIONS OF THE GASTRONOMIC FIELD

Taken as the systematic, socially valorized pursuit of culinary creativity, gastronomy began with the 19th century, and it began in France. The very term came into public view in 1801,⁶ followed by *gastronome* two years later to designate a new social status of the consumer of elaborately prepared fine food.⁷ Like any new social practice, gastronomy drew on a

⁶ The standard reference is to the quite dreadful poem of 1801, "La Gastronomie, ou l'Homme des champs à table" (Gastronomy, or the man of the fields at table) by Joseph de Berchoux. But since a term appears in print well after it has been in circulation, Berchoux is undoubtedly more of a scribe than an inventor. What is important is the role that gastronomy and its derivatives came to play in very short order to designate a practice that was perceived as new. Following Elias [1939] 1994, pt. I) in taking words as sociocultural indicators and seeing dictionaries as repositories of convention, I note that although the French have officially been gourmands since the 14th century and modern gourmets since the 18th, not until the 19th did they become gastronomes. *Gastronomie* first appeared in French in 1623 as a translation from the Greek of the (lost) poem of the Epicurean philosopher and culinary sage, Archestratus (Athenaeus 1969, 1:445–47, 2:237). Here, as elsewhere, all translations are my own unless otherwise indicated.

⁷ The *gastronome* was invariably male. Beyond the fact that men held the purse strings and haute cuisine was a very expensive pursuit, the public culinary sphere was inhospitable to women.

nexus of social, economic, and cultural conditions. It shaped to its own ends the standard exemplar of cultural communication linking supplies, producers, and consumers in a set of common understandings. For gastronomy, this model translates into: first, abundant, various, and readily available foodstuffs; second, a cadre of experienced producers (chefs) in a culturally specific site (the restaurant), both of which are supported by knowledgeable, affluent consumers (diners); and third, a secular cultural (culinary) tradition.⁸ All of these elements—the food, the people and places, the attitudes and ideas—came together in early 19th-century France with a force hitherto unknown and, indeed, unsuspected.

Foods

Paris has long been known for its profusion of foodstuffs and range of food providers—from butchers and caterers to pastry makers and cabaret owners—all of which prompted an appreciative Venetian ambassador in 1577 to report that “Paris has in abundance everything that can be desired.” With food coming “from every country . . . everything seems to fall from heaven” (cited in Revel 1979, pp. 150–51). Two centuries later, the great urban ethnographer Louis-Sébastien Mercier showed a city even more intensely involved in satisfying the gustatory needs and desires of its inhabitants with an estimated 1,200 cooks at diners’ beck and call ([1788] 1994, 1:1011). Even so, the gastronomic level of 19th-century Paris was unmistakably of a different order, fueled, as it was, by more and more wealthy people as well as more and more varied foods brought faster from further away.

In Europe as a whole, the 18th century saw the end of the cyclical famines that had regularly ravaged the continent for centuries and had been such a part of everyday life. In response to demographic pressures, production increased as the expansion of the transportation system trans-

pitable to women: chefs as well as gastronomes were male. The host whose duties Grimod de la Reynière spelled out with such care could only be male. Moreover, as with other urban spaces (shops, parks, public transport, and above all, the street), its inherent promiscuity gave the restaurant an uncertain moral status that effectively excluded upper- and middle-class women. At the most extreme, the gastronome dined alone. Certain of these writings are both misanthropic and misogynist. See, for one example, the “Discours d’un vrai gourmand: Avantages de la bonne chère sur les femmes” (Discourse of a true gourmand: Advantages of good food over women) (Grimod de la Reynière 1984, 2 128–34).

⁸ Cf. the preconditions identified by Chang (1977) and Freeman (1977) for the appearance of a Chinese cuisine in the 12th century. Although these analyses are framed in terms of a culinary *product*—Chinese cuisine—like most commentary, they construct “cuisine” comprehensively, as a set of structured culinary *practices* and *texts* uniting producer and consumer, which I term “gastronomy.”

formed agriculture from a subsistence to a commercial enterprise geared to an increasingly broad market (Teuteberg and Flandrin 1997, pp. 725–26). Specifically for France, with the end of the food shortages of the immediate revolutionary period and despite the British naval blockade, the early century proved a period of alimentary abundance, certainly for the urban elites responsible for making gastronomy a distinctive social practice. The great chef, Antonin Carême was especially sensitive to the deleterious effects of the “great revolutionary torment” on the “progress of our [culinary] art” for 10 years or so and breathed an audible sigh of relief over the far more favorable conditions in the following years ([1815] 1841, 1:xxxii). As observers of the urban scene never tired of pointing out, every country now had its national foods in Paris, with the result that the adventurous diner could take a trip around the world without leaving the table (Briffault 1846, pp. 180–81). When Brillat-Savarin observed with evident pride that a Parisian meal could easily be a “cosmopolitan whole” ([1826] 1839, p. 329), this acute observer of culinary mores meant what he said. In support of the claim that foods came from all over, the 16th-century visitor to Paris gave a list of the French provinces; 19th-century claimants were talking instead about Europe, Africa, America, and Asia.

Restaurants: Producers and Consumers in Public Space

The haute cuisine of the ancien régime served the court and the Parisian aristocracy, but modern culinary creativity centered in the restaurant. Although the restaurant antedated 1789—the first urban establishment by that name dates from 1765—the Revolution set the restaurant on its modern course of development. By doing away with all restrictions on which establishments could serve what foods in what form, the abolition of the guilds spurred culinary competition and prompted a number of former chefs to the now-exiled members of the aristocracy to put their culinary talents in the service of a general elite public (as opposed to a private patron). The restaurants they opened became a notable feature of the urban landscape. Finally, the demise of the monarchy and the court ended the partition of political, commercial, and cultural life between Versailles and Paris, which henceforth was concentrated entirely in the capital. Politicians and businessmen, journalists, writers, and artists flocked to the city and to its restaurants. It was not simply the dramatic increase in population—Paris doubled in size between 1800 and 1850—that was so important a condition for the gastronomic field. The fluid population of largely middle-class transients moving in and out and around the city stimulated the development of eating establishments of many sorts; the hundred or so restaurants found in Paris in the late 18th century increased by a factor of six during the empire and by the 1820s numbered over 3,000

(Pitte 1997, p. 773). As Brillat-Savarin recognized at the time, competition became intense once it became clear that "a single well prepared stew could make its inventor's fortune." As a result, self-interest "fired every imagination and set every cook to work" (1839, p. 324).

That competition was vertical as well as horizontal; the range of restaurants—the consequent economic, social, and culinary stratification that they solidified—was as important a factor in setting up the gastronomic field as the production of haute cuisine in a select few of them. For the observer who regarded gastronomy as the one social force left untouched by "successive upheavals of civilization" (*Code gourmand* 1827, preface), there were several others who were clear that it was part and parcel of a new regime—political, social, and economic. Gourmandise, like elections, had moved from the "summits" of society to its "lowest classes," with the result that the social division that really counted in contemporary France was the one drawn between cooks and diners (Périgord 1825, p. 12). The restaurateurs who made their market niche further down the culinary scale may not have been numbered among the "artists" and the "heroes of gastronomy," but their contribution to the social order was seen as every bit as important (Brillat-Savarin 1839, pp. 324–26).⁹

The diners who rushed to the Parisian "temples of gastronomy" in the new century were certainly as affluent as the aristocrats who had sustained the haute cuisine of the ancien régime. But the gastronomic elite of the early 19th century was socially far more mixed than their predecessors and, to judge by the directives of the nascent culinary journalism, markedly more insecure. Like the renowned connoisseur and consummate courtier-politician Talleyrand, some carried over their *savoir vivre* from the old to the new regime;¹⁰ others, no doubt the majority, had the wealth

⁹ The connection between gastronomy and suffrage became something of a cliché, both phenomena taken as signs of modern times and of the democratization of French society. It is worth noting that both practices remained confined to elites. At the time these works were written, during the Bourbon Restoration (1815–30), the right to vote was determined by the amount and kind of taxes paid and enfranchised approximately 1% of the adult male population. The July Monarchy (1830–48) expanded the voting base to some 8% of the population without altering the basic system. But, just as political life actively involved many nonvoters, notably impecunious scholars and intellectuals, so too gastronomy touched a public that did not dine in the great restaurants (see the discussion of Balzac, below). In both cases, the striking development of publishing and journalism was a primary factor in this general cultural diffusion. There were, of course, other populations—untouched by either restaurants or the texts that talked about them—about whom few culinary journalists talked. Briffault (1846, chap. 5) stands out with even his minimal attention to hunger and to "People Who Do Not Dine," from unemployed workers to prisoners.

¹⁰ The political opportunism of Charles-Maurice de Talleyrand-Périgord (1754–1838) was only slightly more notorious than his love of fine food, allegedly leading a political opponent to remark that the only master Talleyrand never betrayed was Brie cheese

but sorely lacked the *savoir faire*. These were the opulent arrivistes addressed by the culinary journalist Alexandre Balthazar Laurent Grimod de la Reynière, whose detailed instructions on how to be a proper guest as well as a correct host aimed at translating the aristocratic culinary culture of the *ancien régime* for the use of a new public ([1808] 1983). However, the public restaurant, not the private gathering, was the primary vehicle institutionalizing gastronomy as a social and cultural practice in early 19th-century France. Even the eating societies that served as important points of culinary encounter met in restaurants where gastronomy was not simply on view but open to all comers. By relocating culinary creativity and fine dining from private homes into public space, the restaurant offered an ideal, semipublic venue for the display and affirmation of status in a bounded space that simultaneously defined nondiners as non-elite and marked all diners as members of the elite.¹¹ In this manner, competition among diners drove the competition among restaurants.

The participatory disposition of the restaurant contrasted sharply with the imposing banquet spectacles of the *ancien régime*, where the king dined alone in full view of the court. The differences between these two culinary modes are by no means trivial. Whereas the banquet makes use of elaborate, often multitiered culinary creations to manipulate space in the service of a communal spectacle, the restaurant regulates time to effect intimacy. The individual dish offers little scope for the spectacular creation, but it favors the singularization of presentation. The public setting depended upon not only new culinary standards but also a different conception of the meal. Courses were no longer served French style, *à la française*—where many different dishes for a single course are laid out on the table at the same time—but in the simpler, modern style, *à la russe* (it was the supposed innovation of the Russian ambassador during the 1810–11 season)—where a single dish is served for each course to all diners. Against the dramatic display of the traditional French service that so clearly suited the hierarchical arrangement of the *ancien régime* banquet—where one's place at table largely determined which dish one actually ate—its adaptability to variable numbers of individual diners made service *à la russe* a perfect system for the restaurant. Despite the persistence of elaborate banquets for ceremonial occasions like the paradigmatic-

(Guy 1962, p. 119). Carême, who was Talleyrand's chef for a time, categorically affirmed his gastronomic supremacy (1833, 1:xiii–xvi).

¹¹ Boundary-setting mechanisms identified by Goblot (1967) as defining for the French bourgeoisie include the *baccalauréat* degree and its various means (Latin references, e.g.) and the sober formal dress for elite men. Dining out in elegant restaurants was another sign with which the elite distinguished itself from nonelites and upheld group solidarity. More generally on culinary stratification, see Bourdieu (1984, pp. 177–99).

cally extravagant dinner in 1900 for more than 22,000 French mayors (over twice the number of soldiers served by Carême at a similar occasion during the Restoration), the select restaurant gradually relegated the banquet to an even more exceptional, and evidently ceremonial, occurrence. By virtue of this placement of gastronomy in the public sphere, the restaurant anchored the gastronomic field in a fixed institutional basis.

A Secular Culinary Tradition

The final piece in the foundation of the gastronomic field was a secular culinary tradition that offered both institutions and individuals a common intellectual base. Culinary conceptual autonomy presupposed the consideration of food for its own sake and the ideological subordination of religious, symbolic, or medical concerns to the gustatory, however imperfect the separation of the culinary from the symbolic and the medicinal might actually be (Flandrin 1997b). For even though religious interdictions and directives are fundamental to a great many cuisines, they do not in themselves constitute a cuisine. There is no Jewish or Christian or Muslim cuisine; there are, instead, many culinary traditions that negotiate dietary restrictions and ambient cultures and agricultures to construct a given set of culinary practices. For the largely Catholic consumers in France, the process of liberation from the extraterrestrial and the extragustatory entailed diverting attention from the negative associations of gustatory pleasure. First among these was gluttony (*gourmandise*), classified by the Church as one of the seven deadly sins.¹² Criticism from secular quarters invoked sobriety as well as the physical and moral health of both individuals and the social order. In 18th-century France the *Encyclopédie ou Dictionnaire raisonné des sciences des arts et des métiers* ([1751–80] 1966) joined gluttony to a second deadly sin in its definition of *cuisine* as “the lust for good food” (4:537) and *gourmandise* as the “refined and disordered love of good food” (7:754). Such “experiments in sensuality” denature food, which is transformed into “flattering poisons” that “destroy one’s constitution and shorten life” (4:537–39). Following a tradition found in writings as divergent as the Old Testament, Plato, and Herodotus, the authors of the *Encyclopédie* articles conjured up lurid descriptions of the excesses of the late Greeks and decadent Romans to make the point that any thing or practice that reaches beyond nature is not only useless but noxious and

¹² The other six are avarice, anger, envy, pride, lust, and sloth (*avarice, colère, envie, orgueil, luxure, paresse*). The seven sins were codified in the 6th century as deadly or, as French has it, “capital” sins, since these dispositions (rather than acts) were at “the head of,” and therefore responsible for, a multiplicity of sinful acts (anger, for example, leading to murder).

is as destructive of political character as of individual integrity. Gourmandise, it ruled, is considered a merit in countries "where luxury and vanity reign [and] . . . vices are elevated as virtues" (7:753-54). From an individual sin, gourmandise became a social vice, its spread in society a conspicuous sign of the flagrant corruption of the body politic.

Fortunately for the development of French cuisine, these negative judgments of delectable pleasures were offset by strong countervailing pressures from the monarchy and the court, and it was these demands that set the course of fine French cooking. In France, as at many other European courts, public dining rituals elaborated spectacular displays of status and power that reinforced attachment to ruler and court through the manipulation of social distance and spatial proximity (Elias 1983, chap. 3; Wheaton 1983, chap. 7). When Voltaire equated superfluity with necessity in his poem, "Le Mondain" (The man of the world, 1736), and defined excess as utility ("anything superfluous—a real necessity"), he spoke to and for this elite around the court, which indulged in phenomenal luxuriance in many domains—precisely the milieus against which the *Encyclopédie* inveighed so zealously. Even cookbooks joined these culinary polemics. In 1739, the *Lettre d'un pâtissier anglois au Nouveau Cuisinier Francois* (Letter of an English pastry maker to the new French cook) took a stand against the arid intellectualism and the frenetic luxury that took food and consumers away from the primal simplicity of the past. The very next year saw the response: *L'Apologie des Modernes ou réponses du Cuisinier Francois, . . . à un pâtissier anglois* (Defense of the moderns or responses of the French cook), which applies to cuisine the Enlightenment discourse on science (Girard 1977, p. 519; Flandrin, Hyman, and Hyman 1983; Hyman and Hyman 1997).

This discord between indulgence and restraint had not been resolved a century later. However, the terms of the debate had altered considerably, and they had done so because gastronomy had changed the rules of the game. For a privileged witness, we may take the monumental dictionary-encyclopedia of Pierre Larousse (1866-79): *gourmand* and *gourmandise* received two columns of discussion; *gastronomie* was allotted more than four. More significant still was the moral positioning of gastronomy. Next to *gastrolâtrie* ("the passion for good food pushed to a sort of cult . . . incompatible with generosity" [8:1065]) and *gastromanie* ("love of good food pushed to excess" [8:1066-67]), *gastronomie* came across as a model of discipline, control, and moderation. Even though *gourmand* retained many of its pejorative implications, designating an individual "who eats eagerly and to excess," there were no negative connotations for the *gastronome*, who "loves, [and] . . . knows how to appreciate good food." Consequently, "No one blushes to be a gastronome any more," Larousse declared with his habitual assurance, "but at no price would one

want to pass for a gourmand or a drunk" (8:1397). Because gastronomy was both a science and an art, the gastronome could even be considered something of a philosopher-diner, the antithesis, in any case, of the unreflective eater whose lack of self-control led to the gluttony reproved by the Church and castigated by the *Encyclopédie*: "The gourmand only knows how to ingest; the gastronome moves from effects to causes, analyzes, discusses, searches, pursues the useful and the agreeable, the beautiful and the good." This modern construction of culinary fervor sloughed off negative connotations onto the gourmand, the glutton (*glouton* or *goinfre*), or the *gastrolâtre* (someone who "makes a god of his stomach"), thereby securing for the gastronome the lofty position of high priest for this new cult.¹¹ It is important to recognize that this new distinction between the gourmand and the gastronome, the glutton and the connoisseur, had everything to do with an articulated knowledge that every keen observer would recognize. But such powers of observation could be sharpened only through close acquaintance with the emerging body of culinary texts.

The encyclopedia commentary of Larousse suggests the sea change in attitudes about culinary consumption that made it possible to assume, rather than argue, the social significance of gastronomy, an assumption that was itself a prime indicator of "field effects"—the behavioral and expressive repercussions of a cultural practice that had achieved the independence imputable to a field. Yet, neither the social and cultural conditions that generated the new institutions and beliefs nor the new actors and sites devoted to the culinary arts sufficed to create a cultural field. The endogenous standards and models of authority that allow social reproduction, the multiplying subfields and interlocking networks that define the positions and determine what is at stake in the field, the social presence that originates in links to other cultural fields—these field characteristics assume and depend upon a different kind of support.

To move cuisine out of the kitchen and off the dining table—that is, to carry culinary practices from the sphere of immediate material produc-

¹¹ The paradigmatic recasting of vice as virtue and redefining sin as socially useful comes in a novel by the immensely popular writer Eugène Sue. *Gourmandise* is one in the series of *The Seven Deadly Sins*, which was written before and during the Revolution of 1848. Sue's demonstration of the social utility of all the deadly sins accords gluttony (*gourmandise*) the pivotal role. All eight nephews and nieces of the hero are engaged in food production (pastry maker, fish monger, grocer, bread maker, game supplier, butcher, wine merchant, and captain of a merchant vessel with the emblematic name of *Gastronome*, who imports foodstuffs from the colonies). The reformist socialist author set up a profit sharing scheme, not unlike those proposed in other of his novels, and assembled all the "sinful" and "sinning" protagonists of the first six novels for a joyous repast at the end of *Gourmandise*, conviviality reinforcing the positive functions of this erstwhile social sin.

tion and consumption into the broader cultural arena—compelled diffusion, and diffusion demanded a cultural product of a different order. The inherently ephemeral, irremediably private nature of the material culinary product places severe limits on the cultural currency of the culinary arts; food, after all, must be destroyed to be consumed, and, in purely alimentary terms, consumption is strictly individual. Because the material product itself cannot be diffused, culinary practices and products must have an intellectual form that can be put into general cultural circulation. To the extent that cuisine depends on oral transmission, its status as a general cultural artifact and practice remains precarious. The words and texts of an expansive culinary discourse, not the dishes and meals of a circumscribed and confined culinary practice, fixed the culinary product and gave it an existence beyond the sphere of immediate culinary production. Accordingly, the gastronomic field is structured by the distinction between the material product—the foodstuff, the dish, or the meal—and the critical, intellectual, or aesthetic by-products that discuss, review, and debate the original product.¹⁴ The relentless intellectuality of the one is as necessary to the gastronomic field as the insistent materiality of the other. In a paradigm of what cooking is all about, culinary discourse transformed the material into the intellectual, the imaginative, the symbolic, and the aesthetic. The cultural construct that we know today as French cuisine was largely the accomplishment of this discourse, and it was this discourse, secured in texts, that consolidated the gastronomic field.

CULINARY DISCOURSE

Culinary discourse did not, of course, originate either in the 19th century or in France. Western Europe can boast of cookbooks from the 14th century, as well as all kinds of writings concerned with food—literary works, scientific and medical treatises, ethnographic observations, and scholarly ventures, such as the translations of the mammoth work by the Greek

¹⁴ This distinction situates cuisine at the opposite end of the production-criticism continuum from literature—where the original product (the literary work) and the critical interpretation make use of the same vehicle—words. In this respect, cuisine, like music, is a performative art. As such, it depends on words for its social survival—recipes (scores) that make it possible to reproduce the original (Of course, one can question whether the original product is the recipe on the page or the dish on the table.) On this continuum, the plastic arts lie somewhere between the literary and the performing arts, because, although there is a disjuncture between the medium of creation-production and the idiom of criticism, there is a tangible, more or less durable product. The overlap of (creative) writers and critics has a good deal to do with the intellectuality possible in writing. From this point of view, cuisine represents the intellectualization of a sensual, material product—food.

culinary reporter-sage, Athenaeus (1557 and 1612 into Latin; 1680 and 1789 into French). In culinary affairs as in so many others, the 17th century is the turning point. Although France earlier lagged behind other European countries in the production of cookbooks, beginning in the mid-17th century a spate of cookbooks thrust cuisine into the public arena and set off the first episode in the debate replayed by every generation since over the merits of "old" (in this instance, largely medieval culinary practices) versus the "new" and "modern."

Clearly, the 19th century built upon the legacy of the ancien régime. Just as clearly, it was a new era. The new element in early 19th-century France was the particular configuration of culinary discourse, the multiplication of culinary genres, and the sheer volume of apposite writing, with the whole very much a function of the rapidly expanding publishing and journalistic market. The paradigmatic texts ranged from Grimod de la Reynière's *Almanach des Gourmands* (The gourmands' almanac) in 1803 to Carême's summum three decades later, *L'Art de la cuisine française au dix-neuvième siècle* (The art of French cuisine in the 19th century). Although writing anchors every cultural field, the transitory nature of culinary products renders the gastronomic field absolutely dependent on a textual base. For language allows sharing what is at once the most assertively individual and yet, arguably, the most dramatically social of our acts—eating. If words turned food into culinary texts, these texts inserted gastronomy into a field. They set the culinary agenda and instituted the cultural debates that defined the gastronomic field as well as the logic that determined relations within this field.

Five genres of gastronomic writing laid the foundations for the gastronomic field. The "professional" genres that contributed most obviously and directly to the formation of the gastronomic field were: the gastronomic journalism of Alexandre Balthazar Laurent Grimod de la Reynière (1758–1838); the culinary treatises of Antonin Carême (1784–1833); and the cultural commentary and protosociology of Anthelme Brillat-Savarin (1755–1826). These three authors were witnesses to, even as they were agents of, the modernizing gastronomic society. Each aimed to systematize culinary knowledge; each contributed to the formalization and, hence, the very definition of modern French cuisine. But the professionals did not operate alone. Indeed, it would seem to be characteristic of cultural fields generally, as it certainly is of the gastronomic field in early 19th-century France, that the larger social impact of the field is importantly a function of extraprofessional participation and noninstrumental writings. The greater the association of nonspecialists, the more numerous the connections to other cultural fields and to society at large and the greater the social impact. It is precisely this kind of indirect participation that points to the fourth and fifth genres of gastronomic writings, the political philoso-

phy of Charles Fourier (1772–1837) and the novels of Honoré de Balzac (1799–1850).¹⁵ These five exemplary works, taken together, hold a key to the understanding of the place occupied by the gastronomic field among French cultural fields and within French society. Defining gastronomy as a practice and establishing the genres of culinary writing, professional and nonprofessional writings together established the textual archive of the gastronomic field in France.

Grimod de la Reynière

Grimod de la Reynière was the first gastronomic journalist. Known before the Revolution for outrageous culinary extravagances, he made use of his extensive knowledge to enlighten 19th-century elites, most of whom he thought woefully ignorant of the most elementary gastronomic practices. Grimod put the culinary conscience of the ancien régime at the service of the new in an era of acute social and culinary change. His assessments of culinary establishments (restaurants, food suppliers, stores, and products) and practices ([1803–12] 1984, vol. 8) aimed at ordering a culinary world turned topsy-turvy (1983, p. xxxi; 1978, pp. 311–16). The new gastronome was not to the manners born but a self-made man in need of instruction. It was Grimod who assumed responsibility for the culinary “catechism” (the term he used for his *Manuel des Amphytrions* [Manual for hosts], 1808). If the fine art of carving had been lost in the “revolutionary torment,” Grimod would rectify the situation. He argued that a host who did not know how to carve was every bit as shameful as an owner of a magnificent library who did not know how to read (1983, pp. xxxiii–xxxv, 3). In a proposal that would be taken up by many others in various forms and working from the assumption that gastronomic science had advanced by leaps and bounds, Grimod advocated the creation of still more culinary institutions, including professorships in the lycées (1983, p. xxxdi), various

¹⁵ It is not by chance that gastronomy owes its existence to founding fathers, and this despite the “natural,” “logical,” and traditionally dominant associations of women with food and feeding. But those associations concern the domestic order, whereas gastronomy occupied the public domain. Furthermore, in statements similar to those made relative to artistic activity, women were deemed incapable of culinary creativity. They were “slaves to routine,” as a not atypical remark at the time asserted. If a man “hadn’t grabbed a hold of the frying pan, [culinary] art would have stayed where it was, and we would still be eating Esau’s lentils and Homer’s roastbeef” (Périgord 1825, p. 121). Carême, for his part, characterized modern cuisine as both “virile and elegant” (1841, dedication, 1xxviii). On the whole, the professionalization of the culinary arts over the 19th century excluded women in the most public, upper reaches, an exclusion that is still apparent in upscale French-oriented restaurants today, in France and elsewhere (see Mennell 1985, pp. 201–4; Cooper 1998; Ferguson and Zukin 1998).

gastronomic societies, and an elaborate system of what he called "legitimations"—whereby product samples were sent for evaluation to "tasting juries" composed of "professors in the art of Gourmandise" (1984, 1:xvii, 2:xix–xx, 3:xxx, 4:vi–vii). Grimod's gastronomic enterprise was an immediate and enormous success, with 22,000 copies of the *Almanach* sold in several editions over the four years following publication. "Thanks to the progress of knowledge and philosophy," Grimod declared in the *Journal des Gourmands et des Belles* of 1806, undoubtedly thinking of his own contributions, "gourmandise . . . has become an art" (1806, 1:23).

Carême

Grimod de la Reynière laid down the law for consumers and dealt exclusively with the "theory" of gastronomy—he bragged that he had never put on an apron, and, in any event, only individuals having the misfortune to live outside Paris stood in any need of recipes (1806, p. 107). Antonin Carême, on the other hand, legislated for the professional practitioner.¹⁶ Acknowledged by contemporaries and later generations alike as what he proclaimed himself to be, namely, the founder of modern French cuisine, Carême had an immense influence. His systematic examination of the bases of French cuisine, first, of pastry in all its permutations (which covered savories along with sweets), and then all of French cuisine, from soups to sauces via fish, stuffings, bouillons, roasts, and game, provided the methodological basis for the subsequent expansion of the profession later in the century. Unlike Grimod de la Reynière, who overtly took his cues from an ancien régime that he viewed with considerable nostalgia, Carême disdainfully dismissed the old cuisine in favor of the new—"19th-century French cuisine will remain the model for culinary art" (1833, vol. 2, pt. 4, p. 13). All extant cookbooks and treatises on cooking—"these sorry books" ([1822] 1842, 1:5)—were to be jettisoned in favor of his totally original synthesis; his was "the honor and the merit" of giving "our great cuisine" the treatment that it deserved, and what is more, he had not "borrowed anything from anybody" ([1828] 1986, p. 20). Cuisine aspired

¹⁶ As he repeatedly stressed, Carême's works were not simply collections of recipes but culinary treatises. But they are much more than that, especially as Carême very much fancied himself a writer, even while lamenting his untutored style (1841, 1:xi). Not only is virtually every recipe preceded and followed by observations, anecdotes, and sundry remarks, each volume contains inordinately ambitious disquisitions of a more general order, a lengthy "Preliminary Discourse," a "Parallel of Old and Modern Cuisine," a "Philosophical History of Cuisine from the discovery of fire to the present," "Aphorisms, Thoughts, and Maxims," "How Napoleon Ate in Exile on St. Helena," "A Critical Review of the Grand Balls of 1811–1812," and similar commentary.

to the status of both a science and an art, and Carême's goal was to turn those gastronomic aspirations into culinary practice.¹⁷

Carême very explicitly addressed a modern, newly expanded public, which his works were designed to expand further still. He rested his case for the general utility of his work on the fact that even women could profit by the volumes to instruct their cooks at home (the directive gives an idea of the bourgeois nature of that public). Accordingly, he simplified the meal—four courses for a formal dinner instead of the usual eight—and he pared down the banquets from those of his ancien régime predecessors, giving more space per person and placing fewer and smaller serving platters on the table, and so on (cf. the foldouts of table setups; 1842, vol. 2). Ever mindful of expense for this broader audience, Carême defended himself against accusations of excess, coming back again and again to the practicality of his cuisine. Even the early works on pastry made a great point that these semiarchitectural confections were “easy to make,” however implausible it may seem to any nonprofessional reader. His great valedictory treatise began with a discussion of the humble beef stew (*pot au feu*), disdained by 18th-century and modern authors alike even though, as Carême carefully pointed out, it furnished the principal source of nutrients for the working classes. The *pot au feu* is, moreover, where the sauces, and therefore modern French cuisine, begin. It was this “19th-century spirit of analysis” (1833, 1:lxvi) that led him to a chemical analysis of what actually happens when the housewife puts the stew pot on the fire (1833, 1:3–4).

Yet Carême also harked back to an earlier era. He made his reputation working in the houses of the great, not in restaurants, and was very proud of his relationships with his illustrious patrons (notably, Talleyrand, Tsar Alexander I of Russia, the Prince Regent of Britain, and Baron Rothschild), which earned him the sobriquet “the chef of kings, and the king of chefs.” For Carême as for Grimod de la Reynière, the ideal gastronomic couple was the gastronome and the chef: “The man born to wealth

¹⁷ Although French cookbooks had for some time urged cooks to forswear the strong spices of medieval cooking (ginger, coriander, cinnamon, allspice) (Flandrin, Hyman, and Hyman 1983), it was Carême who most decisively and emphatically replaced these with herbs (thyme, basil, savory, bay leaf, parsley, chervil, tarragon) and plants (garlic, shallots, onions) (Carême 1833, 1:lx-lxiv). A description by an English visitor to the Rothschild household where Carême was in charge gives an idea of this new, simplified yet complex, and self-consciously modern cuisine as it appeared on the table: “Its character was that it was in season, . . . up to its time, . . . in the spirit of the age, . . . no trace of the wisdom of our ancestors in a single dish; no high-spiced sauces, no dark brown gravies, no flavour of cayenne and allspice, no tincture of catsup and walnut pickle, no visible agency of those vulgar elements of cooking, of the good old times, fire and water. . . . Every meat presented its own natural aroma, every vegetable its own shade of verdure” (Morgan 1831, 2:415–16).

lives to eat, and supports the art of the chef." Correspondingly, he had only contempt for the "rich miser [who] eats to live" (1833, 2:v, vi-vii). The great era of gastronomy was not the period he was writing for in the 1820s but the empire in the first decade of the century, with its great patrons and its opulent "extras"—extraordinary, spectacular banquets, including one for 10,000 soldiers under tents along the Champs-Élysées—the menus of which he gives in nostalgic detail (1842, chap. 13). Not surprisingly, he was highly critical of the modern service à la russe that works so well in restaurants and is so ill-suited to banquets: "Our French service is more elegant and more sumptuous . . . the very model of culinary art" (1842, 2:151).

Carême's dual culinary allegiance meant first of all that he had an extraordinarily sharp sense of his own worth, in the kitchen and on the printed page. He went on at length about the long nights spent experimenting with different dishes, the hours in the Royal (then Imperial, then Royal) Library researching culinary achievements from earlier eras, the killing hours buried in coal-burning kitchens to the great detriment of his health, the expenditure of his own moneys in the service of gastronomy, the lack of culinary appreciation on the part of the French in contrast to foreigners, and so on. Many of these statements sound like nothing so much as the ideal-typical romantic creator lamenting his sacrifices in the name of art or science. Yet, this exalted sense of self went along with, and was set against, an equally strong sense of placement vis-à-vis the culinary past, its present, and its future. A prodigious knowledge of culinary traditions grounded Carême's insistence on his place as the creator of modern French cuisine. Virtually all of his works contain more or less extensive comparisons between traditional and modern culinary methods, and one even presents a "Philosophical History of Cuisine" (1833). All the works explicitly addressed his *confrères*, his practicing culinary contemporaries, of whom three receive book dedications.¹⁸ Carême was ever mindful of training subsequent generations of chefs; in the work that was the culmination of his career, *L'Art de la cuisine française au dix-neuvième siècle*, a section of "Remarks and Observations about the Young" advised young

¹⁸ Carême's game was rather more complicated because the chefs and steward (*maître d'hôtel*) to whom he dedicated his works were carefully situated with respect to their elite employers: M. Mueller (*Le Pâtissier pittoresque*) was the chief steward for the Russian Tsar; M. Boucher (*Le Pâtissier royal*), the steward for Talleyrand's household; the Robert brothers (*Le Maître d'hôtel français*), chefs who had worked in the most illustrious houses in Paris and Europe. Not until *L'Art de la cuisine française au dix-neuvième siècle* (1833) did Carême, who remained extremely conscious of his humble origins, feel secure enough to dedicate his work directly to his patron (and even then the dedication was at one remove, since the dedicatee was Madame Rothschild).

chefs who must make their reputations in culinarily less favored times than the glory days of the empire.

Carême envisaged a systematically organized profession, with "cook-offs" and examinations for the best chef, the best dish, and so forth. Here again, Carême bridged the old system and the new. These culinary competitions were aimed at regulating market competition. First, the imposition of professional standards adapted to market conditions the guarantee of artisanal quality once afforded by the strict regulations of the guilds. Further, by publicizing both professional standards and the winning professionals, these instances of culinary certification made a strong bid to control the market. As any winner of such an award today will testify, this kind of highly public professional certification possesses significant market value. With such institutional mechanisms connecting practitioners around a common set of experiences, techniques, and values and to the public, Carême's contribution was essential to the formation of the gastronomic field. The subsequent development of professional cooking in the latter half of the 19th century would not have been possible without his example and his works, the rules he laid down, the techniques he explicated, and the ideal of the creative chef that he embodied.

Brillat-Savarin

From Grimod de la Reynière and Carême to Brillat-Savarin, gastronomy was converted from a practice and a technique to a topic of general discussion and analysis beyond those directly concerned with material production and consumption. For although Brillat-Savarin wrote about food, it was, precisely, all the talk about food by a wonderful witty conversationalist that immediately made his *Physiology of Taste* (1826) the totemic gastronomic text that it remains today. To the culinary paradigm of chef-diner, Brillat-Savarin added the reader, the consumer for whom the cultural, rather than the material product, is the primary concern. Unlike the journalist, who addresses customers, clients, and diners, or the chef, who targets practitioners of the culinary arts, the commentator-analyst reaches to these indirect consumers—the readers whose culinary consumption is indirect because it is noninstrumental.¹⁹ The kind of culinary

¹⁹ Instrumentality can be (re)defined by the reader's circumstances. Gillet (1993) argues that cookbooks constitute true gastronomic literature and that recipes not only can but should be read as an exercise in literary gourmandise. Changes in culinary as well as reading conditions similarly affect instrumentality. Medieval recipes, e.g., mean little to us today, even if we could procure the exact ingredients, while, at the other end of the spectrum, the complexity and technicity of the professional cooking Carême initiated make it all but impossible for nonprofessionals to give his recipes anything but a noninstrumental reading. And what reading should we give to a work

commentary practiced by Brillat-Savarin and generations of his disciples places gastronomy within the larger intellectual and social universe. For Grimod de la Reynière and Carême, the culinary text was chiefly instrumental, a means to the primary end of producing or consuming what anthropologists term the "food event," that is, the dish or the meal. For Brillat-Savarin, the text was its own end, a status hardly altered by the few recipes included in the work. The often-noted stylistic qualities of the *Physiology of Taste*—the anecdotal mode, the witty tone, the language play—give this work an almost palpable literary aura. Even so, and however important this style for the immediate and continuing success of this work, the decisive literary connections were determined less by individual achievement than the nature of the work, which places the *Physiology of Taste* with more clearly literary-intellectual genres within the larger culinary discourse.

More decidedly than Grimod de la Reynière or even Carême, Brillat-Savarin conceived of gastronomy as a distinctly modern social practice. His admission that a fear of falling behind the times had prompted him to undertake the study illustrates the degree to which he equated gastronomy with modernity, an intellectual enterprise representative of a contemporary body of knowledge and a nontraditional, analytical attitude toward food. The many anecdotes and witticisms should not obscure the claims this work made to theoretical, historical, and even scientific understanding. The subtitle—*Meditations on Transcendent Gastronomy—an Up-to-Date Theoretical and Historical Work*—confirms that these bonds were not incidental and singular but structural and generic. Brillat-Savarin's assertion was doubly encyclopedic since gastronomy is both comprehensive and foundational, drawing as it does on the natural sciences—physics, chemistry, physiology—and on learning of every sort, including cuisine, commerce, political economy, and medicine. The youngest science was born when the chemist, the scholar, and the political economist took cuisine out of the kitchen into the laboratory and the library. Never again could food be confused with either a sin or a mere bodily function. Like Grimod de la Reynière and Carême, Brillat-Savarin took considerable pains to distinguish gourmandise from gluttony. It was "the passionate, reasoned and habitual preference for objects that flatter taste," just as gastronomy was "the reasoned knowledge of everything that concerns man and nourishment." A mental activity dealing with the senses, gastronomy relied on refined sensuality but even more on intelligence: "Animals fill themselves; man eats; the intelligent man alone knows how to eat" (Brillat-Savarin 1839, pp. 63, 317, 65).

like *The Alice B. Toklas Cook Book*? Should the celebrated hashish fudge really be attempted?

The second component of Brillat-Savarin's analysis is properly sociological. In effect, the *Physiology of Taste* suggested a model for a sociology of taste as this 18th-century philosophe metamorphosed into a sociologist, subjecting the "pleasure of the table" to its ever-changing social contexts.²⁰ Brillat-Savarin elevated gastronomy to the rank of a science and justified what he called "social gourmandise" by its exceptional social utility. The *Physiology of Taste* offers a few menus from sumptuous and more modest meals, describes some dishes, and even gives some recipes. But Brillat-Savarin indulged comparatively little in the vice that plagues food writing of every era, namely, the rehashing of gustatory memories. It was not simply the variety of French cuisine that marked the *Physiology of Taste* as a sociological enterprise but the correlations that Brillat-Savarin established between the social and culinary attributes of taste. Gastronomy, as Brillat-Savarin argued the case, was the science that explored those relations. Clearly, as well, it was a social science that examined even as it joined a more general discourse on class and class distinctions. Taste taken in this very specific sense became another powerful marker of class as gastronomy came to define individual appetite in collective terms.

The science of gastronomy had such a formidable task because the social diffusion of gourmandise did not eliminate so much as it complicated the gastronomic hierarchy. If there are individuals whom nature has "predestined" to be gourmands and whose very physiognomy betrays their predilection, more interesting to Brillat-Savarin are those figures destined to assess social positions and professions in terms of their penchant for gourmandise: financiers, doctors, men of letters, and the pious (*les dévots*) head the list (1839, pp. 176–85). Brillat-Savarin further devised a series of "gastronomic tests" calibrated to income, with menus ranging from five courses for a 5,000-franc income, six courses including truffled turkey for diners in the 15,000-franc income bracket, to the nine courses of extravagant, complex dishes appropriate for those with an income of 30,000 francs and over (1839, pp. 188–89). A necessary but not sufficient factor, money could not be construed as in any way a cause of taste preferences. It was, rather, an indicator of "gastronomic class," which intersected with social class, to be sure, although not nearly so neatly as it had under the ancien régime. Brillat-Savarin prudently avoided ranking the financier's fare "better." In keeping with the neutralizing language of science, the evalua-

²⁰ The *Physiology of Taste* seems to have been the first work to exploit *physiologie* as a sociological as opposed to biological or medical concept. In the decade that followed, the *physiologie* came to refer to a short essay of 4–10 pages published in a small and usually illustrated volume, which purported to identify the characteristic social types, institutions, or accoutrements and behavior of modern society (Ferguson 1994, p. 82–90).

tions of the the *Physiology of Taste* were those of the "dynamometer," which registered increasing force as one ascended the social ladder. The dishes able to test the gastronomic faculties of the stolid bourgeois *rentier* would not be regarded as worthy of examination by the "select few" invited by a banker or governmental minister. With its own hierarchy and its variable standards, the world of gastronomy reproduced the contradictions and the ambiguities of postrevolutionary society—Brillat-Savarin noted a dinner that assembled gourmands of the fourth (highest) class (1839, p. 193).

To be sure, the *Physiology of Taste* did not take full account of that public world. Brillat-Savarin found the restaurant, which his "philosophical" history of cuisine placed among the "latest refinements," so new in the 1820s and so different an institution that no one had thought about it enough. With only one "Meditation" devoted to "Restaurateurs," Brillat-Savarin did not really take up the challenge. Yet he was both disconcerted and intrigued. On the one hand, the restaurant was an element of democratization; on the other, restaurants catering to solitary diners fostered an excessive individualism possibly destructive of the social fabric (1839, pp. 318–24). In any event, and however significant a social phenomenon the restaurant might be, the private gathering supplied the model of sociability for the *Physiology of Taste*. Grimod de la Reynière's work had already made it clear that this world too had been greatly altered by the increased circulation of individuals and their culinary habits, the availability of goods and services, and the culinary pluralism that the restaurant represents. By virtue of its interpretation of cuisine as a collective enterprise, defined by the consumers rather than the producers and governed by the social as opposed to the alimentary situation, the *Physiology of Taste* placed cuisine squarely in the public domain. In Brillat-Savarin's work, the science of human nourishment—gastronomy—became something more—a science of society.

The Nationalization of French Cuisine

The texts of this gastronomy also served as a vehicle for a distinctively French culinary nationalism. If French and foreigner alike have long considered cuisine quintessentially French, the explanation lies importantly in an expansionist culinary discourse that relentlessly associates (good) food and France, and has done so for some three or four centuries. From the 16th century onward, European cuisines generally were moving in their separate directions, away from the commonalities of earlier culinary modes (Flandrin 1997a). The Venetian ambassador in 1577 was not alone in remarking on the singularity of French foodways. Then too, many of the cookbooks that began to appear in the mid-17th century made a point

of the "Frenchness" of their cuisine. But in virtually all of these instances, "French" was not a geographical but a social reference, and French cuisine was French by virtue of the court and the aristocracy. Cookbook writers invariably vaunted their elite connections—in titles like *Les Soupers de la cour* (Suppers at court; 1755) and in flowery dedications to noble patrons. Works expressly destined for the bourgeoisie, like the *Cuisinier royal et bourgeois* (The royal and bourgeois cook; 1691) or the phenomenally successful *La Cuisinière bourgeoise* (The [female] bourgeois cook; 1746), proposed simplifications of the courtly models (Mennell 1985, pp. 80–83), but those models remained in force. This aristocratic model also supplied the basis for the first truly international cuisine, carried by French chefs and cookbooks to court kitchens throughout Europe. Just as European aristocracy in the 17th and 18th centuries spoke and wrote French, so too it "ate French" and relegated native culinary traditions to the status of poor relations.²¹ The culinary writers of 19th-century France found themselves in the enviable, and unique, position of working within a celebrated indigenous culinary tradition. The increasing centralization of French society and the attendant concentration of French cultural institutions—conditions matched nowhere else in Europe—further reinforced these associations between the nation and elite cuisine. Of course, France also had nonelite, regional cuisines, but other countries had only such cuisines (which is why Carême, like many of his ancien régime predecessors, was called to cook abroad).

Like the other nationalisms that flourished in the 19th century, French culinary nationalism drew on texts. That "French cuisine" was itself the product of texts has led some to deny the very possibility of a national cuisine. Because, for these critics, a cuisine is product based, it can only be local (Mintz 1996, chap. 7), which makes "French cuisine" little more than an intellectualized Parisian artifice totally dependent upon "true," regional cuisines.²² Such statements tend to confuse the plurality of culi-

²¹ Mennell (1985, chap. 5, pp. 102–33) argues that the distinctive culinary practices of the English gentry and prosperous farmers, which had no equivalent in France, were "decapitated" when the urban aristocracy adopted French culinary models beginning in the late 17th century. French professional cuisine developed from this courtly and later urban model; English professionals, including more women, evolved out of the more domestic culinary practices associated with more modest households and the countryside.

²² Jean-François Revel (1979, chap. 8) similarly denies the possibility of a national cuisine but for different reasons. The opposite of essentially conservative, traditional regional cuisines is an "international" cuisine defined by techniques and methods and rooted in the search for originality. While these categories make sense from a strictly culinary point of view, they elide the culinary nationalism that identifies a particular culinary configuration as "French." On the complex process by which fries ("french fries") became the "alimentary sign of Frenchness" and the sequence that transformed

nary practices with the critical coherence of a culinary tradition. As a formalized set of culinary practices, any cuisine is necessarily the product of a culinary discourse, and, to reach beyond the confines of the originating group, that discourse needs texts. Even regional cuisines relied on texts to be constituted as such; as self-consciously formulated culinary codes, regional cuisines in France were themselves products of culinary writings as early as *Le Cuisinier gascon* (The gascony cook; 1740) but diffused more widely beginning, once again, in the 19th century (Cservo 1997). The textual reliance of a national French cuisine finds confirmation in the rise of national cuisines in late 19th-century Italy (Camporesi [1989] 1993) and 20th-century India (Appadurai 1988) where the absence of a prestigious indigenous culinary model ensured the dominance of regional or foreign identifications until countered by a significant text or set of texts.²¹

The gastronomical writings of Grimod de la Reynière, Côté, and Brillat-Savarin, along with those of their many critics and disciples, gave the requisite textual basis to a specific set of culinary practices. These writings became, as they had not been theretofore, identified with the country as a whole. The generalization of interest in gastronomy, attested in many quarters and acknowledged by all three authors, supplied a public eager for such materials. These three "professionals"—the journalist, the chef, the sociologist—invested cuisine in France with a force, a value, and a presence that have long been considered characteristic of French foodways. In effect, the culinary discourse in which they engaged "nationalized" French cuisine, and it did so by imposing the standards and arrogating the authority that would henceforth define positions in the gastronomic field.

The authoritative tone of so many of these writers has much to do with the continual jockeying for position in the emergent gastronomic field. Culinary discourse, in fact, offers a perfect illustration of symbolic violence, nowhere more vividly or with greater impact than in France. The very term "gastronomy"—from *gastro* (stomach) and *nomos* (law)—signaled the importance of rules and regulations in this culinary construction. (Côté made much of standardizing the orthography for menus [1841, 1:lvii–lxxxix].) There was no brooking dissent with Grimod de la Reynière

Camembert cheese from a clearly regional product into one recognized by the French themselves as "French," see respectively, Barthes ([1957] 1972) and Boisard (1992).

²¹ For Italian cuisine, the text was Pellegrino Artusi's *La Scienza in cucina et l'Arte di mangiar bene—Manuale pratico per le famiglie* (The art of eating well; 1891), which, in Camporesi's argument (1993, pp. 113–52), created a code of national identification that did more for national unification than Manzoni's great epic novel, *I promessi sposi* (The betrothed). That certain dishes (corn polenta, potato gnocchi, spaghetti with tomato sauce) came to enjoy a national status was due importantly to Artusi's textual promotion.

or Carême or any number of their epigones. To be sure, recipes, most particularly written ones, must be authoritative; at the very least, every recipe implies a culinary authority. But recipes were not alone in adopting a legislative mode. Other components of French culinary discourse had just as many, if different, pretensions to authority. Witness the great proliferation of texts that supplied the codes by which gastronomy was regulated and lived. The *Code Gourmand* of 1827 trumpeted itself as a *Complete Manual of Gastronomy, containing the laws, rules, applications and examples of the art of living well*; it was followed by *The French Gastronomer, or the art of living well*, written by "the former Authors of the *Journal des Gourmands*" (among whom was Grimod de la Reynière). This work was in turn succeeded by the *Perpetual Almanach of Gourmands, containing Le Code gourmand with its applications, rules and meditations of transcendent gastronomy* (the reprise of Brillat's subtitle making its particular bid for authority).

The consciously authoritative tone of such works is as unmistakable as it is significant. Although Brillat-Savarin published the *Physiology of Taste* anonymously, the title page proclaims the author's status as a "Professor" and "Member of several learned societies." Analysis in this work is inextricably tied to prescription. The off-hand, humorous tone of the work belies the imperious manner of the precepts and principles. The dichotomy was intentional; the use of "I" and "me," Brillat-Savarin explained, supposes a "confabulation" with the reader, who is free to "examine, discuss, and even laugh." But when he comes armed with the "redoubtable *we*," it is a lecture, and the reader must submit. "I am, Sir, oracle," the authorial persona warned, citing *The Merchant of Venice* in a comical yet firm directive. "And, when I open my lips, let no dog bark!" (1839, p. 36). So too, Carême passed over no opportunity to cite his right to readers' consideration, the many and spectacular successes, the eminent personages for whom he had worked, the breadth of his experience and extent of his research, and the utility of his innovations and method.²⁴

All commentators agreed that the culinary arts were French, so much so that it is not too much to speak of a culinary nationalism. Grimod de la Reynière announced that he would not regret all the care (and heart-

²⁴ That authority might be contested. The gastronome and the chef were touted as the ideal culinary couple, but the relationship was fraught with tension from competing demands due most particularly to the chef's ambiguous and fairly untenable position as simultaneously an artist and an artisan. Each of these roles assumes a different relationship with a patron/consumer/client. Cf. a contemporary "defense" of gastronomes against the "aberrations of [an innovative artist's] delirious imagination" that railed against the absence of a "culinary law" that could "contain the culinary art within its true boundaries . . . and would put . . . a brake on the propagation of doctrines pernicious to gastronomy" (Périgord 1825, p. 4).

burn) that went into the *Almanach des Gourmands* if "the national glory in all aspects of alimentary art" progressed because of it (1984, 2:xx-xxi). Carême simply pronounced France "the motherland of anyone who entertains guests; its cuisine and wines are the triumph of gastronomy and it is the only country for good food" (1986, p. i; 1833, 2:i). In an earlier work, he boasted that "this absolutely new Treatise . . . will give new luster to our national cuisine" (1841, dedication), and in another, he boldly asserted that he had already refuted all the "ridiculous books that are a disgrace to our great national cuisine" (1842, 1:5). The work he had in mind, quite unfairly given the date of its publication, was the cookbook that first associated France with a cuisine—*Le Cuisinier françois* of 1651. The oxymoronic pretensions of a contemporary's cookbook entitled the *Universal English Chef* had Carême so beside himself that he proposed a cook-off, it being a foregone conclusion that French chefs would win any culinary contest (1841, 1:xvi).

Not that French cuisine lived in splendid isolation. Indeed, its "genius" lay in the strength of its capacity to assimilate foreign elements. Although French cuisine was "indisputably the first in the world," as Grimod de la Reynière recalled in 1806, it could become richer still by drawing on foreign foods. But appropriation also meant transformation, which is to say, "Frenchification." French cuisine would incorporate exotic foods "by perfecting them" (1984, 3:295 n. 1). Some 20 years later, Brillat-Savarin had recourse to this same notion of acclimation to illustrate the relationship of French cuisine to foods, seasonings, and drinks "of foreign preparation" (1839, pp. 316-17). For many, "French" was equated with "classic." The publication of English and German cookbooks in the 1820s sent a conservative culinary critic into great (mock) diatribes against the "Romantic" cuisine that paid no attention to the rules of culinary Classicism. The vogue for English literature was one thing in this period of dramatic clashes between Romantics and Classics in print and on the stage, but, for the author of the *New Almanach for Gourmands*, food was a truly serious affair. Protesting vehemently against the wholesale importation of dishes that "at great expense produce a dubious taste and nausea," this commentator enjoined the French to make judicious choices in English cuisine no less than in English literature. If roast beef and boiled potatoes did no harm to (French) gastronomic sensibilities, plum pudding, salt beef, and mutton soups—"bizarre preparations unsuitable for either our health or our climate"—should be left to those whose "ironclad palates" were accustomed to such fare, just as all the garlic-laden dishes from the south of France should remain where they belonged (Périgord 1825, pp. 104-8).

Reminiscent of the "theory" of climates traceable to Montesquieu and, in literary studies, to Mme de Staël, this understanding of cuisine in terms of national tradition and temperamental suitability asserts what was im-

plied in the more serious, more "professional" works by the founding fathers of the gastronomic field: the culinary nationalism, even chauvinism, that would become inextricably allied with French cuisine. The constitution of a gastronomic field depended upon a redefinition of haute cuisine as a national cuisine. To be sure, the haute or grande cuisine of the ancien régime had been considered "French"—witness the claims, which Carême felt obliged to refute, of the 17th-century cookbook, *The French Chef*. Moreover, any number of commentators from the 17th century onward affirmed the intrinsic, virtually "natural" affinities between France and fine food. But the "nationalization" of this cuisine occurred in the 19th century when an overtly class culinary model turned into a national cultural phenomenon. Although the resulting culinary product was still class based, the class had shifted to include an extended gastronomic elite for the consumption of the material product and a still larger gastronomic public for reading culinary texts.

Nowhere were the rules of culinary conduct more highly and more authoritatively codified than in 19th-century France, inspired at least in part by dismay over evident social change. It is hardly surprising that foreign models of cooking should have aroused anxiety about culinary disorder. "Soon," one writer intoned after reporting a spate of works purporting to introduce English and German cooking to the French, "all the aberrations of an art which has no limits and acknowledges no rule will be transplanted here" (Périgord 1825, p. 106). Just so, the author of the *Code Gourmand* called for a "fixed and immutable code" to contain the excesses of the almost libidinous Gastronomy, "the queen of the world" (1827, prologue). Another critic asserted that "cooking, like the drama . . . ought to follow Aristotle's rules" (*Almanach perpétuel des Gourmands contenant Le Code Gourmand, et des applications, règles et méditations de gastronomie transcendante* 1830, p. 106). On the other hand, the assimilationist "genius" ascribed ever more emphatically to French cuisine not only legitimated, it virtually demanded a broad spectrum of sources to enrich native components. However, borrowing would take place strictly on French terms and in accordance with preexisting national norms and precepts.

In this manner, culinary discourse constructed a paradigm for the cultivation of a self-consciously national identity, a wonderful illustration of the cultural work of nation building so characteristic of 19th-century Europe. French cuisine was one more "invented tradition" (Hobsbawm 1983) that shored up a national identity. Cuisine and gastronomy were part of a more general process of cultural nationalization through the imposition on the periphery of the values, the norms, and even the language of the center (Weber 1976), an ideal choice for a new and somewhat shaky republican regime seeking to elide political conflict by promoting emblems to unify the country rather than divide it (Nora 1996–98). Assimilationist

by nature, French cuisine subsumed all the others in a perfect synecdoche of France. If regional products and dishes continued to play a vital role in national culinary construction, they did so as subordinate parts of an incomparably greater whole.²⁵

The vehicle of this synecdoche, the link between part and whole, between region and country, is found in the expansive culinary discourse of the 19th century. This connection explains why the ties between the literary and the gastronomical are not accidental but a constitutive feature of the gastronomic field. Cuisine in France became fully French by virtue of the discourse that incorporated food practice into a socioculinary tradition and a gastronomic code. The culinary discourse elaborated in early 19th-century France created a substantially new, and almost infinitely extendible, set of gastronomic consumers—readers. The French cuisine secured in these many gastronomical writings could be consumed far more readily on the page than at the table. Because capacities for intellectual ingestion far exceed the physiological capacities of the most voracious appetite and most determined glutton, the public for culinary texts is potentially immense. It was this public, which overlapped to a variable degree with actual diners, that provided the most solid foundation for the gastronomic field, determined the positions staked out in that field, and permitted, even demanded, discussion, commentary, and criticism. But allowing repeated “consumption” of the same meal or food, writing and reading neutralize the orality of food. Just as writing fixes speech, so too it stabilizes food. Culinary discourse controls consumption, which it transforms into an intellectual activity. Virtually all the definitions of gastronomy stress this intellectual displacement, often citing Brillat-Savarin’s definition of gastronomy as the “reasoned knowledge” of everything to do with what we eat (1839, p. 65).

²⁵ The synecdochal perspective, the perennial dialogue between center and periphery, between haute cuisine and popular cuisines, between intellectualized cuisine and product-based regional cuisines, is revealed with particular clarity in the introduction to a novel that introduced the paradigmatic gastronome, Dodin-Bouffant (conceived as a homage to Brillat-Savarin). The author admits hesitating to publish a work on an apparently frivolous topic so soon after the immense suffering sustained in the Great War of 1914–18. But should he neglect one of the “oldest and most essential of French traditions?” “A quiche lorraine . . . or a Marseillaise bouillabaisse . . . or a potato gratin from Savoy has all the refined richness of France, all its spirit and wit, its gaiety . . . the seriousness hidden beneath its charm, . . . its malice and its gravity, . . . the full soul of its fertile, cultivated rich earth, of which its aromatic cream sauces, snowy poultry, delicate vegetables, juicy fruits, savory beef and frank, supple and ardent wines, are the blessed manifestations” (Ruoff [1924] 1994, pp. 12–13). (Ruoff also coauthored a massive, multivolume culinary history of the French provinces.) On Proust’s hymn to French culinary sensibility in *À la recherche du temps perdu*, see Ferguson (in press).

Fourier

Culinary discourse also explains the gastronomic associations specific to a modernizing French society. Beyond the texts directly concerned with the culinary production and consumption—the works of Grimod de la Reynière, Carême, and Brillat-Savarin—are those that analyze and dramatize food as a total social phenomenon. The philosophical writings of Fourier and the novels of Balzac offer excellent examples. More emphatically than any other text at the time, and by its very utopianism, the social order imagined by the utopian philosopher Charles Fourier (1772–1837) demonstrates how food works as a total social phenomenon to mold institutions no less than individual behavior.²⁶ Fourier based his philosophical system on the social utility of pleasure, specifically, the principle of attraction, the two most powerful principles of which are sex and food, or in his terms, love and gourmandise. No more than sex was gourmandise an individual matter. Fourier constructed an entire social system to turn these individual pleasures to social account.²⁷

Yet even as Fourier set his work apart from the ambient gastronomic discourse, he invoked many of the same themes and principles—the scientific nature of gastronomy, the importance of culinary judges and juries, the crucial distinction between gastronomy and gluttony, the social utility of gastronomy in a time of rising economic prosperity. In every case, Fourier took the precept outside the contemporary social order (which he derivatively referred to as “Civilization,” in contradistinction to the projected social order that he called “Harmony”). Gastronomes, writers as well as practitioners, and even the best among them like Grimod, produced nothing better than “gastro-asinities” (*gastro-âneries*) (1966–68, 6:255, n. 1). Even Fourier’s cousin Brillat-Savarin was no better than any other so-

²⁶ The irregular publishing history of Fourier’s work makes his contribution to culinary discourse more conjectural than for the others. Although his first work appeared in 1808 (*Théorie des quatre mouvements et des destinées générales*), his last remained in manuscript until 1967 (*Le Nouveau monde amoureux*). But Fourier’s ideas were known well before the (fragmentary) edition brought out by his disciples in the mid-1840s (Beecher 1986).

²⁷ In the vast majority of gastronomic writing, the parallels with sexual activity are irresistible and seldom resisted. It is not by chance that collectivities so assiduously regulate the one and the other to keep the direct sensuality of the individual from disrupting the social order. Second-order consumption also looms large for both. In this as in other domains, Fourier’s writing is conspicuous for its mixture of tones and genres, all of which add up to what can be fairly characterized as controlled delirium designed to convey the fundamental attraction of a new social order predicated on neither justice nor equality but happiness. “The events resulting from this Order will give you, not the objects of your desires, but a happiness infinitely superior to all your desires” (1966–68, 1:170).

called gastronome ignorant of the higher or combined gastronomy, which Fourier baptized *gastrosophy* (8:283) and in which he found “a profound and sublime theory of social equilibrium” (4:130), “the principal mechanism of the equilibrium of the passions” (6:258), and one of the two primary bases of the new social order. The higher gastronomy occupied such an important place in Harmony because it presided over the development of the senses rather than their repression. Fourier premised his entire system on material abundance, which alone could guarantee spiritual abundance (e.g., happiness) (1:77). Contemporary society was based on differential scarcity, whereas the increased production of harmony would spread abundance throughout society. (Fourier broke fractionally with Saint-Simon on just this issue.) The emphasis on plenty led Fourier to define his new culinary order against gastronomy as civilization understood the practice. The moderation preached in 19th-century gastronomic circles was anathema: “A hundred thousand philosophers eat only to keep their passions under control.” His view that moderation is a “travesty of nature” (6:255–56) led him to place all activities in Harmony under the twin signs of profusion and the absence of moderation: prodigious appetites will necessitate five meals plus two snacks a day, men will be seven feet tall, easy digestion will make children strong, and life expectancy will be 144 (1:180, n. 1).

Seldom have the culinary and the social order been more explicitly or more visibly tied and of greater moment than in Fourier’s vision. Like Grimod, Brillat, and Carême, Fourier worked to transcend the gross materiality of food, and, like them too, he was maniacally concerned with detail. But where they defined gastronomy in terms of art and science, he made it the stuff of economics, philosophy, and politics. His gastronomical political economy endowed the proverbial land of milk and honey with an elaborate, complex social organization grounded in a visionary social science. What makes this culinary utopia more than a curiosity is what it reveals about the emergent gastronomic field in 19th-century France. More than any other culinary text, Fourier’s writings intellectualized gastronomy, and they did so by making connections to established intellectual enterprises of unimpeachable legitimacy—philosophy and political science, or what in 19th-century France were called the “moral and political sciences.” Today, we would certainly add sociology, like gastronomy an intellectual innovation of the early 19th century.

Balzac

As Fourier carved out a place for philosophy and the social sciences in the gastronomic field, Honoré de Balzac (1799–1850) set up the literary

relations.²³ Balzac built on a long tradition of literary culinary commentary that, in French literature alone, dated from at least Rabelais and Montaigne in the 16th century. The undoing by unbridled appetite(s) had been standard comic fare from the Greeks (cf. Athenaeus 1969). Balzac's perspective differed from that of his predecessors because it was resolutely and self-consciously modern. He boasted of being the "Secretary" to French society (1976–81, 1:11), and certainly one of the striking contributions of Balzac's novels—one greatly appreciated by Marx and lavishly praised by Engels—is the dramatic ethnography of the nascent industrial capitalism of postrevolutionary France. Balzac used the realist novel to define contemporary French society, and, in that definition, food and feeding loomed large. Balzac's characters eat a good deal, and, more important still, Balzac attaches great significance to the consumption of food as a social and psychological indicator. What visitor to Paris does not identify with Lucien de Rubempré, when he decides to "initiate himself in the pleasures of Paris" at a restaurant where a single dinner eats up the 50 francs that would have lasted him a month at home in the provinces. Small wonder that the impecunious young man soon finds his way to the Latin Quarter and a menu at 18 *sous* (1976–81, 5:271, 292, 294–95). Balzac understood, as Grimod, Côté, and Brillat-Savarin did not, the significance of the restaurant as a privileged location of gastronomic and other modernity—a semipublic, semiprivate urban space of dubious moral and variable culinary quality. He also showed the degree to which the restaurant acted as something of a safety valve for the increasingly regulated mores of bourgeois society (Aron [1973] 1975).

Dinners and diners turn up all over Balzac's work. But cuisine is more than a strong marker of distinction. In Balzac's last, and darkest, novel, *Cousin Pons* (1846), gastronomy became the stuff of tragedy, a "bourgeois tragedy," as Balzac called the genre in another setting, "with neither poison or dagger or blood but . . . crueler than all the calamities in the house of Atreus" (1976–81, 3:1148). This novel made gastronomy an actor like money, an agent of the dramas, particularly of the tragedies, on which Balzac constructed his modern society. Gastrolatry is the good-hearted, hapless Pons's tragic flaw, his sin gourmandise. Balzac modernized that sin and brought the punishment up to date. Pons is not set upon by divine retribution but by identifiable social forces, specifically, avaricious relatives who defraud him of an incomparable collection of antiques. Like the fine collection of antiques that he has amassed with such loving care,

²³ A generation younger than the other founding gastronomic fathers, Balzac knew their work well. His gastronomic credentials include a *Gastronomic Physiology* (1830), the entry on Brillat-Savarin in the *Biographie Michaud* (1835), and the *New Theory of Lunch* (1830) (1938, 2:43–47, 62–63, 671–76).

Pons's worship of fine food expresses his fine artistic nature and offers compensation for personal disappointments.²⁹

In giving gourmandise tragic dimensions, Balzac broke with a philosophical tradition that restricted expression of the baser senses—touch, smell, and especially taste—to base-born characters and to the baser genre of comedy. Balzac's reinterpretation of the literary mode for such expression expanded and strengthened the connections between literature and gastronomy and, hence, between the gastronomic and the literary fields. Even as Balzac was positioning himself and his work in the emerging literary field, he was also, if less obviously and less consciously, defining a position in the nascent gastronomic field. The same work addressed both fields, albeit differently. In the one—the literary field toward which Balzac directed his strategies—the literary work supplied the primary cultural product, and reading that work constituted the primary cultural consumption. In the other—the gastronomic field—those same novels were part of a second-order consumption. Reading is an indirect culinary practice, and reading literary or intellectual works is the most indirect of all, furthest on the continuum from the instrumentality of cookbooks and similarly directive texts.

Straddling Fields

Such intersections between continuously shifting cultural fields are not only possible, they are all but inevitable. A given individual may be central to one field and peripheral to another, may change subjective stance toward and objective position in one but not the other, and so on. Other things being equal, the more fields in which an individual occupies a position, and the more central those positions, the more symbolic power or capital at that person's disposal. A "cultural career" is made of the trajectories followed over a lifetime. Exploiting the opportunities offered by this sort of "cultural mobility" is what Balzac did so magnificently. The literary field was not simply the sociological context in which he occupied a position and in which he maneuvered. It also furnished him with a subject: Balzac was not only an actor in the nascent literary field of 19th-century

²⁹ "For him celibacy was less a preference than a necessity. Gourmandise, the sin of virtuous monks, opened her arms to him, and he threw himself into them as he had thrown himself into the adoration of art. . . . For him good food and Bric-a-Brac were substitutes for a woman" (1976–81, 7:495). Balzac continues, reproving Brillat-Savarin for not placing enough importance on "the real pleasure" to be had at table. "Digestion, by using human forces, constitutes an inner battle which, for gastrolaters, is the equivalent of the greatest climaxes (*jouissances*) of love." Even more completely than his collection, Pons's gourmandise satisfies the desire for the total merger with the desired object.

France; this author was also, and in some of his greatest works, an ethnographer of that field.

However, as the metaphor implies, straddling fields is possible only to the degree that the fields themselves intersect. Balzac's position in, and analysis of, the gastronomic field was a function of just such a field intersection. Although it is possible to occupy positions in noncontiguous, even unconnected fields, a high degree of intersection tends to be characteristic of cultural fields. Such convergence is especially notable in modern French society, where the long-term concentration of cultural institutions and activities in Paris favors the intersection of cultural fields as well as the interaction of individuals. Field intersection more generally is promoted by the common educational training received in the elite secondary school, which is another factor behind the conspicuous ties between the literary and political fields in France (Clark 1979).

The associations between the gastronomic field and the literary field are vital to the situation of cuisine among French cultural products and to the position of the gastronomic field in the hierarchy of cultural fields. Although the second-order consumption of this "literary gastronomy" places it on the outer reaches of the gastronomic field, this textual consumption is a crucial element in the diffusion of the values and the traditions that govern the field. The literary work effects the ultimate transcendence of the material and transformation of the sensual. This work of transformation points to the decisive distinction between cuisine and gastronomy and their respective functions in the gastronomic field. Cuisine, or culinary codes, concerns production; its injunctions are largely instrumental, its practice more or less site-specific. Gastronomy, on the other hand, is a code that pertains to consumption; it is grounded in primarily gratuitous, that is, noninstrumental, discourse. Each of these cultural products operates within the gastronomic field; each is a necessary component of the cultural consciousness characteristic of that field. For the gastronomic field to come into existence, cuisine had to connect with gastronomy, and culinary production had to be linked to culinary diffusion. It did so through texts that also made connections to other cultural fields—literature especially prominent among them.

OTHER CULTURES, OTHER FIELDS

Considering gastronomy as a cultural field brings a number of long-standing questions into sharper focus. The development of the gastronomic field from the mid-19th century in France points to the processes that moved the field from beginnings to consolidation. The subsequent professionalization of cooking beginning in the 1880s worked off the continued expansion of restaurants, particularly in the great hotels catering to the nascent

tourist industry that acted as important training grounds for cadres of French chefs; at the end of a long career that began in the 1870s, the highly influential Auguste Escoffier boasted that he had sent some 2,000 French chefs from his kitchens all over the world (1985, p. 193). The automobile turned increasing numbers of diners into culinary tourists, and, beginning in the 1930s, the Michelin restaurant ratings established a national geography that at the same time set up a hierarchy of French cuisine. Professionalization brought further proliferation of texts, specialized journals, newspapers, and reviews, which address the domestic cook, the professional chef, and also, increasingly, readers for whom cooking is akin to a spectator sport. New media such as radio and television (and latterly the Internet) are today integral parts of, and active actors in, the gastronomic field.

The evolution of the gastronomic field in France impels us to reconsider the supposed disappearance of cultural singularity in an increasingly global economy. For foodways in France, the "McDonaldization" of food production (by no means entirely imputable to American corporations) and the continuing changes effected by the European Economic Community have raised fears of a possible loss, or at the very least a significant weakening, of distinctive cultural identities. Such fears are by no means new—witness the negative reception given German and English cookbooks in the 1820s or a 1924 warning that "France would no longer be France" when a French meal was no different from repasts elsewhere (Ruoff 1994, p. 13).³⁰ But the extent and strength of the field, with its extensive organization, its range of institutions, the values and beliefs those institutions perpetuate, and the self-consciousness that characterizes the field as a whole, lead us to posit the cultural field as a site of resistance to the (real or perceived) eradication of cultural difference. The more tightly organized the field and the greater its reach, the greater its autonomy and ability to perpetuate its core values. At the same time, the broader cultural resonance of the field, and therefore its capacity for cul-

³⁰ Rick Fantasia (1995) demonstrates that such fears are not without foundation in the France of the late 20th century, given the inroads of fast food and, more important, changes in the eating patterns of the French (fewer women at home to prepare a full-scale lunch, commuting distances that make in-house canteens more practicable for employee as well as employer, different, "modern" foodways more attuned to culinary pluralism). Perhaps the most striking of his findings is the clear distinction drawn by the adolescents interviewed between fast food and traditional forms of restauration, notably the *café*, distinctions based on the very different types of sociability seen as appropriate to each: where fast food represents a rejection of the traditional French culinary norms of the adult world and a quick "American" fix, the *café* remains important as a place to drink and to talk much as it has been for a century and more. Fischler (1990, pp. 212–17; 1997) places these debates within an international context, noting that loss of cultural identity is a concern in Spain and Italy as well as France.

tural resistance, is importantly dependent upon connections to other cultural fields and other institutions. In other words, a cultural field owes its singular position to a particular configuration relating the part to the whole, the field to the larger society, in which the larger social ties both temper and enhance the autonomy of the field. Complete independence of a field from its larger context can make no sense; it would be the cultural equivalent of solitary confinement. Isolation would nullify any larger impact as surely and as effectively as direct control by institutions.

Besides the great advantage of locating French culinary practices in a specific socioculinary setting, taking gastronomy as a field makes it possible to identify what is French about food in France. Although it may be commonly agreed, and not by the French alone, that gastronomy is somehow "innately" French, only in the 19th century, however important the ancien régime contribution, can one identify anything resembling a national culinary discourse. Of course, assigning gastronomy to the French "character" or unique geography or exceptional climate begs the question, all the more so since most of the institutions, ideologies, and practices that express these character traits originated, again, only in the 19th century. For, if culinary creativity in France was highly visible in the ancien régime and elite consumption singularly conspicuous, the gastronomic field, like a number of other cultural fields, arose in postrevolutionary France. These fields defined and were defined by publics that were larger, more expansive, and more heterogeneous than their prerevolutionary counterparts. In fixing these culinary practices in a circumscribed space, the gastronomic field allows us to distinguish between what is distinctively French and what is more generally modern about these culinary practices, what French cuisine and French culinary practices share with other cuisines, as well as the elements that set French foodways apart.

The concept of the gastronomic field allows us, for example, to make better sense of the connections between French and Chinese cuisine. The high degree of codification of the rules governing both culinary production and consumption clarifies, and justifies, the comparison and sets both against the regional cuisines in each country. Chinese and French elite cuisines build on strong, prestigious elite culinary traditions originally tied to a central government and an urban elite. Both, as well, were sustained by a significant textuality (Freeman 1977), and, in China as in France, visible cultural enthusiasm is tempered, channeled, and contained by authoritative culinary and gastronomic codes. Aside from the actual techniques of preparation and cooking (which are, indeed, very different), Chinese cuisine differs most importantly from the French in its evident philosophical overlay (Chang 1977). By contrast, French cuisine emerged out of a resolutely "secular" environment. In more recent times, against the continued support of the French government for various culinary ini-

tatives—the *Chambre syndicale de la haute cuisine française*, the *École nationale*, and *Centre national des arts culinaires* founded in 1985, the Web sites, the classes that teach very young school children how to taste, the culinary competitions such as the *Meilleur Ouvrier de France* for different categories of cuisine, the commercial manifestations such as the *Salon du chocolat*, and so on—the policies of the Communist regime that did so much to destroy elite institutions in China greatly affected culinary practices by interrupting the course of culinary tradition and thwarting the practice of gastronomy (cf. Wenfu 1988). Such close, direct political control, even if less stringent than in the recent past, makes it unlikely that a gastronomic field in the full sense of the term could be identified in present-day China.

Gastronomic Fields, Culinary Cultures, and Restaurant Worlds

If China is often compared to France for the refinement and complexity of its cuisine, the United States is more likely to be invoked as its polar opposite. And, although culinary America is a very different and substantially more sophisticated place than it was only a few decades ago, it does not offer the culinary unity or authority requisite for a gastronomic field. There is, in the first place, no cultural product on which to base a field because there is no American cuisine, that is, no culinary configuration identified with the country as a whole.²¹ The foodways of Colonial America were either unwritten (Native American), foreign (Dutch, English, or Spanish, depending on settlement patterns) or both (divers African). Strong regional identities yield more or less local, product-based regional cuisines—New England, *Tex-Mex*, Southern, Cajun. Because these cuisines tend to be identified by dishes (North Carolina versus Texas barbecue, New England versus Manhattan clam chowder), they are susceptible to great variation (chowders alone would take us on a tour around the country). More recently, this distinctive American pluralism has come to include the foodways of newer immigrant groups, a number of which fast-food chains have made an integral part of the American diet (Belasco 1987; Mintz 1996); in *Pizza Hut* and *Taco Bell*, *McDonald's*, *Burger King*, and *Kentucky Fried Chicken*, foreigners (Fantasia 1995) as well as Americans find the most visible common element of American foodways. Whatever other culinary unity Americans may have comes not from food but

²¹ If there is no American cuisine, there is an identifiable American diet (prevalence of fast food eaten out and prepackaged foods eaten at home, high levels of animal protein, salt, fat, and processed sugars and correspondingly low levels of fresh fruit and vegetables, preference for soda over water). See Mintz (1996, pp. 117–22) along with the warnings issued with disquieting regularity by various health authorities

a food event: Thanksgiving. The United States may well be the only nation that harks to a meal as foundational event, that is, one of the founding, and perduring, myths of a singular American destiny (Ferguson 1996). Yet, here too, the legendary meal of turkey, pumpkin, and cranberries gives rise to innumerable variations fixed in regional or ethnic custom (sweet potato casserole with marshmallow topping; spaghetti or chili as a side dish), family tradition (mince pie instead of pumpkin), or idiosyncratic modifications.²² In other words, pluralism wins out even for a food event that is insistently constructed as a defining national occasion.

This cultural pluralism supports, as it is supported by, a relative lack of cultural authority. None of the various national tourist guides (Mobil, AAA) approaches the authority of the *Michelin Guide*, whose annual restaurant ratings in France arouse such great expectation and anxiety on the part of diners and restaurateurs alike. It is symptomatic, and emblematic of American foodways, that the well-received *Zagat* restaurant guides for a number of cities and regions in the United States (and now, Paris) rely on self-selected informants rather than experts. Similarly, in the case of literature, no literary prize awarded in the United States, not the National Book Award, not the Pulitzer, enjoys the authority and the consequent impact on sales of the top literary prizes in France, most notably the Prix Goncourt (Clark 1979; 1987, chap. 1).

Yet, of course, even a society without a cuisine has characteristic foodways, which is to say that it has a culinary culture—a set of identifiable values and representations that have shaped and continue to inform those foodways. A certain degree of (self-)consciousness characterizes contemporary culinary cultures, which is why it can be argued that a French culinary culture emerges prior to the 19th century (Revel 1979; Mennell 1985, chap. 4). What the 19th century added with the establishment of the gastronomic field is the acute consciousness of positions and possibilities for social mobility in a circumscribed social space.

Given that every society has a culinary culture, it falls to the ethnographer to chart that culture and track down indigenous foodways. A culi-

²² The celebrated chef and cookbook writer James Beard, who did so much to promote American foods, nevertheless proposed a heavily gallicized revisionist meal even as he claimed allegiance to the traditional turkey: along with a stern warning against cranberry (it obliterates the taste of the wine), he recommended serving champagne or vodka with a first course of caviar or smoked salmon, followed by a French red wine for the cheese course and another, sweet wine for the pumpkin pie, the whole topped off with kirsch, framboise, or cognac (1965, p. 323). It is not without import for my overall argument that Thanksgiving, as a national food event, is the product of texts, relayed by a panoply of representations: first, the journal of Edward Winslow that recounted the meal of 1621; and subsequently, the proclamation of 1863 by which Abraham Lincoln declared Thanksgiving a national holiday and the annual presidential declarations since.

nary culture is more comprehensive, less concentrated, less necessarily conflictual than a gastronomic field. It is also not centered on a specific cultural product. French culinary culture includes but reaches well beyond French (haute) cuisine and the gastronomic field. Similarly, American culinary culture comprehends much more than the fast food eateries that are so salient a feature of the American landscape. And where texts are essential to the intellectualization of food and therefore the constitution of the gastronomic field, a culinary culture incorporates a wide range of representations, most of which will not be intellectualized or even written—sayings (“Show me a soldier plate”), anecdotes (George Bush’s avowed distaste for broccoli), advertising slogans (“Where’s the beef?”), and images from radio, cinema, and television: Popeye’s association with spinach is indelibly inscribed in the culinary memories of generations of moviegoing Americans.³³ French culinary culture, too, has been shaped by popular sayings (“Dans le cochon, tout est bon” [Everything in the pig is good to eat]) as well as media representations (François Mitterrand’s obituaries made much of the president’s food and dining preferences). Whereas visual images lend important support to the gastronomic field, they are absolutely central to the formation of a culinary culture. By the same token, the texts that play the major role for a gastronomic field are less salient in the more broadly based, less focused, more loosely participatory culinary culture.

If the gastronomic field does not make sense either of or for American culinary culture, what account can be made of the America that dines out more and more, not at McDonald’s but in restaurants situated at the antipodes of industrialized fast food? What about the America that reminds one suspiciously of France, with its adulation of avant-garde chefs and taste for culinary adventure both close to home and in far-flung places? How do we discuss the urban America in which restaurants have been so signal a factor in the reconfiguration of the cityscape and the practices it generates (Zukin 1991, chap. 7)? A production of culture perspective suggests *restaurant world* as an appropriate model, that is, to adapt the technical definition of *art world*, “the network of people whose cooperative activity . . . produces the kind of [culinary] works that [restaurant] world is noted for” (Becker 1982, p. x). Such cooperative networks can exist only in fairly circumscribed social or geographical settings en-

³³ In 1997, the Jewish Museum in New York City presented a wonderful video exhibit of Jewish food on American television from the 1950s to the present (*Chicken Soup and Wontons*, 1997). Similar shows could undoubtedly be mounted for other “ethnic” cuisines of long standing. Movies such as *Eat-Drink Man-Woman* (Chinese cuisine), *The Big Night* (Italian cuisine), and the iconic film of French cuisine, *Babette’s Feast*, are important vehicles of diffusion of elite culinary values and practices.

dowed with mechanisms that promote connection. The sheer size of the United States, the ambient cultural pluralism, the conflicting occupational identifications of chefs and cooks (Fine 1996a) dictate that restaurant worlds in the technical, sociological sense are the exception rather than the rule.³⁴ This restaurant world is structured by a network of high-end restaurants run by self-consciously innovative chef-entrepreneurs. General professional support comes from a number of organizations and periodicals (Dornenburg and Page 1995, pp. 298–304; Cooper 1998, pp. 281–88), but more important for these elite chefs are the elite media representations that diffuse critiques and praise of given restaurants as well as anecdotes about star chefs, who, television at the ready, are likely to turn into media personalities in their own right. Centrifugal economic factors (each restaurant produces a singular cultural product and competes with others in the same market niche) are countered by centripetal social forces generated by close personal and professional connections (Ferguson and Zukin 1998).³⁵ While the density of these elite restaurants is highest in New York city, the network of chefs is nationwide and, not infrequently, international as well.

Each of these models fits with a larger paradigm of assumptions concerning the relationship of food and society. The *restaurant world* focuses on production of a more or less well-defined culinary product—which, in the case of the fin de siècle American restaurant world, can be characterized as avant-garde, eclectic cuisine. A restaurant world coheres through networks of individuals, whereas a *culinary culture* fixed in practices and values is above all a model of culinary reception or consumption. Finally,

³⁴ Cf. the striking absence of such a larger connection in the restaurants studied by Gary Fine (1996b, pp. 133–37) in a medium-size urban setting (Minneapolis–St. Paul). Fine relates these thin networks to the fragmented economic organization of the restaurant industry (each restaurant producing its own, singular product) and the structures of restaurant kitchens. That this world of restaurants is far from a restaurant world is clear from the speculation of a reader Fine cites about what would be necessary to turn the world of these restaurants into an art world (1996b, p. 264). Zukin (1995, chap. 5) confirms this picture in a study of a range of midlevel restaurants in New York City.

³⁵ As concerns patterns of interaction, a number of New York chefs interviewed (Ferguson and Zukin 1998) mentioned that the irregular hours a chef spends on the job effectively restricts socializing to other chefs, both informally (one chef regularly ended the evening in the kitchen of a competitor-friend two blocks away) and formally, through institutions such as the James Beard Foundation and charitable benefits. Thus, simply in the first week of May 1998 as I was revising this article, four of New York's top chefs prepared a reception to benefit the Frick Collection; the James Beard Awards winners were announced at a benefit dinner; and another dinner to benefit the James Beard Foundation brought together 48 chefs from across the country, all of whom had at one time or another cooked in the famed restaurant of the 1980s, the Quilted Giraffe (it closed in 1992).

a *gastronomic field* is structured importantly by a largely textual discourse that continually (re)negotiates the systemic tensions between production and consumption. The model chosen will depend on the theoretical and intellectual agenda. Culinary culture and the restaurant world take us to food; the gastronomic field points us toward other cultural fields and particularly toward the arts.

CONCLUSION

The sociological import of gastronomy or cuisine extends well beyond the particular cultural product. Viewing gastronomy as a field refines our understanding of cultural fields as such, their characteristics and their character, how they operate and evolve, the respects in which they are similar and those where they differ, their connections to the larger field of cultural production and the larger society. From the outset, it has been apparent that the nature, or medium, of the cultural product is a fundamental determinant of the structure of both the field and the relations between production and consumption. Unlike the literary field, where the primary and secondary product rely on the same vehicle—writing—but like the musical field, the gastronomic field is structured not simply by the duality of direct production and indirect, critical commentary—the case for any cultural product—but by the radical disjuncture between the material and the intellectual products and the consequent dependence of the field on the written document. The objection is often raised that words “get in the way” of the primary cultural experience (tasting, seeing, reading, listening) (cf. Wolfe 1975), that their intellectualization neutralizes the senses. But, in cultural fields, there is no getting around words. Whatever the deleterious effects on individual appreciation, it is the discourse of criticism and chronicle that must bear most of the weight of the cultural field.

The premium that second-order discourse places on innovation aligns gastronomy with the arts more generally in modern society where the indirect cultural production sustains the more or less stable configuration that renders conflict creative. Wherever one comes down in the debates over cuisine as an aesthetic phenomenon, it is indisputable that cuisine in France warrants classification among the arts by virtue of the attributes shared with other cultural and specifically artistic fields. The simultaneous susceptibility and resistance to change, the drive toward innovation against the force of tradition, aligns gastronomy with other modern arts that occupy fields that are similarly divided or, more accurately, fragmented among multiple production sites, each of which negotiates invention and convention. Every field will have its distinctive networks and strategies, its bastions of traditionalism along with outposts of innovation.

By simultaneously containing and promoting competition, the field generates the inevitable struggles that are the signs of cultural ferment and creativity. For good or for ill, the arts in contemporary society are tightly bound to their apposite cultural fields. Accordingly, the study of culture will do well to look to the concept of the cultural field to elucidate both the mechanics and the range of cultural experience. The gastronomic field is so useful as a particular example precisely because it enables us to talk more concretely and particularly about cultural experiences that are easily lost to sociological study when language and practice are not aligned in theoretical understanding.

REFERENCES

- Almanach perpétuel des Gourmands contenant Le Code Gourmand, et des applications, règles et méditations de gastronomie transcendante.* 1830. Paris: Barba.
- Appadurai, Arjun. 1988. "How to Make a National Cuisine: Cookbooks in Contemporary India." *Comparative Studies in Society and History* 30:3-24.
- Aron, Jean-Paul. (1973) 1975 *The Art of Eating in France: Manners and Menus in the Nineteenth Century*, translated by N. Rootes. London: Owen.
- Athenaeus. 1969. *The Deipnosophists of Athenaeus of Naucratis*, 7 vols. Translated by C. B. Gulick. Cambridge, Mass.: Harvard University Press, Loeb Classical Library.
- Balzac, Honoré de. 1938. *Oeuvres diverses*, 2 vols. Edited by M. Bouteron and H. Longnon. Paris: Conard.
- . 1976-81. *La Comédie humaine*, 12 vols. Paris: Gallimard-Pléiade.
- Barthes, Roland. (1957) 1972 "Steak and Chips." Pp. 62-64 in *Mythologies*, translated by A. Lavers. New York: Hill & Wang.
- Beard, James. 1965. *Menus for Entertaining*. New York: Delacorte Press.
- Becker, Howard. 1982. *Art Worlds*. Berkeley and Los Angeles: University of California Press.
- Beecher, Jonathan. 1986 *Fourier: The Visionary and His World*. Berkeley and Los Angeles: University of California Press.
- Belasco, Warren J. 1987. "Ethnic Fast Foods: The Corporate Melting Pot." *Food and Foodways* 2 (1): 1-30.
- Boisard, Pierre. 1992. *Le Camembert: Mythe national*. Paris: Calmann-Lévy.
- Bourdieu, Pierre. (1966) 1969 "Intellectual Field and Creative Project." *Social Science Information* 8 (2): 89-119.
- . (1979) 1984. *Distinction: A Social Critique of the Judgement of Taste*, translated by R. Nice. Cambridge, Mass.: Harvard University Press.
- . (1984) 1988. *Homo academicus*, translated by P. Collier. Cambridge: Polity Press.
- . 1993. *The Field of Cultural Production*. New York: Columbia University Press.
- . (1992) 1996 *The Rules of Art: Genesis and Structure of the Literary Field*, translated by S. Emanuel. Stanford, Calif.: Stanford University Press.
- . (1996) 1998. *On Television*, translated by P. Ferguson. New York: New Press.
- Briffault, Eugène. 1846. *Paris à table*. Paris: J. Hetzel.
- Brillat-Savarin, Anthelme. (1826) 1839. *Physiologie du goût ou Méditations de gastronomie transcendente Ouvrage Théorique, Historique, et à l'ordre du jour*. Paris: Charpentier.

- Camporesi, Piero. (1989) 1993. *The Magic Harvest: Food, Folklore and Society*, translated by J. K. Hall. Cambridge: Polity Press.
- Carême, Antonin [Marie-Antoine]. 1833. *L'Art de la cuisine française au dix-neuvième siècle: Traité Élémentaire et Pratique suivi de dissertations culinaires et gastronomiques utiles aux progrès de cet art*, 2 vols. Paris: chez l'auteur.
- . (1815) 1841. *Le Pâtissier royal parisien*, 3d ed 2 vols. Paris: n. p.
- . (1822) 1842. *Le Maître d'hôtel français ou Parallèle de la Cuisine ancienne et moderne selon les quatre saisons*, 2 vols. Paris: J. Renouard.
- . (1828) 1986. *Le Cuisinier parisien, ou L'Art de la cuisine française au dix-neuvième siècle* Paris: Firmin-Didot.
- Chang, K. C. 1977. Introduction to *Food in Chinese Culture. Anthropological and Historical Perspectives*, edited by K. C. Chang. New Haven, Conn.: Yale University Press.
- Chicken Soup and Wontons: Jews and Food on American Television*. 1997. Produced by the Jewish Museum. New York: National Jewish Archive of Broadcasting.
- Clark, Priscilla P. 1979 "Literary Culture in France and the United States." *American Journal of Sociology* 84 (5). 1057–77.
- . 1987 *Literary France: The Making of a Culture*. Berkeley and Los Angeles: University of California Press.
- Code gourmand: Manuel complet de Gastronomie*. 1827. Paris: Ambroise Dupont
- Cooper, Ann. 1998 "A Woman's Place Is in the Kitchen": *The Evolution of Women Chefs*. New York: Van Nostrand Reinhold.
- Csergo, Julia. 1997. "L'Émergence des cuisines régionales." Pp. 823–41 in *Histoire de l'alimentation*, edited by Jean-Louis Flandrin and Massimo Montanari. Paris: Fayard.
- Dorenburg, Andrew, and Karen Page. 1995. *Becoming a Chef*. New York: Van Nostrand Reinhold.
- Elias, Norbert. 1983. *The Court Society*, translated by E. Jephcott. Oxford: Blackwell.
- . (1939) 1994. *The Civilizing Process*, translated by E. Jephcott. Cambridge, Mass.: Blackwell.
- Encyclopédie ou Dictionnaire raisonné des sciences des arts et des métiers*. (1751–80) 1966. Stuttgart: F. Frommann Verlag
- Escoffier, Auguste. 1985. *Souvenirs inédits: 75 ans au service de l'art culinaire*. Marseille: Éditions Jeanne Laffitte.
- Fantasia, Rick. 1995. "Fast Food in France." *Theory and Society* 24:201–43.
- Ferguson, Priscilla Parkhurst. 1994. *Paris as Revolution: Reading the 19th-Century City*. Berkeley and Los Angeles: University of California Press.
- . 1996. "Cuisinel Cuisine! The Memory Machine." Paper presented at the meetings of the Social Science History Association, New Orleans, November.
- . 1997. "Paysages culinaires" Paper presented at the colloquium, "Paysage et identité régionale" Université Stendhal-Grenoble-3, Valence, 15–18 October.
- . 1998. "The Careers of Chefs: 'French' and 'American' Models of Cuisine." Pp. 92–111 in *Eating Culture*, edited by Ron Scapp and Bryan Seltz. Albany: State University Press of New York.
- Fine, Gary Alan. 1996a. "Justifying Work: Occupational Rhetorics as Resources in Restaurant Kitchens" *Administration Science Quarterly* 41:90–115.
- . 1996b. *Kitchens: The Culture of Restaurant Work*. Berkeley and Los Angeles: University of California Press.
- Fischler, Claude. 1990. *L'Homme-civile: Le goût, la cuisine et le corps*. Paris: Odile Jacob.
- . 1997. "La 'Macdonaldisation' des mœurs." Pp. 859–79 in *Histoire de l'alimentation*, edited by Jean-Louis Flandrin and Massimo Montanari. Paris: Fayard
- Flandrin, Jean-Louis. 1997a. "Choix alimentaires et art culinaire (XVI^e–XVIII^e siècles)" Pp. 657–81 in *Histoire de l'alimentation*, edited by Jean-Louis Flandrin and Massimo Montanari. Paris: Fayard.

- . 1997b. "De la diététique à la gastronomie, ou la libération de la gourmandise." Pp. 683–703 in *Histoire de l'alimentation*, edited by Jean-Louis Flandrin and Massimo Montanari. Paris: Fayard.
- Flandrin, Jean-Louis, Philip Hyman, and Mary Hyman. 1983. "La Cuisine dans la littérature de colportage." Pp. 11–107 in *Le Cuisinier français*, edited by Jean-Louis Flandrin, Philip Hyman, and Mary Hyman. Paris: Montalba.
- Fourier, Charles. 1966–68. *Oeuvres complètes*, 12 vols. Paris: Anthropos.
- Freeman, Michael. 1977 "Sung." Pp. 141–76 in *Food in Chinese Culture. Anthropological and Historical Perspectives*, edited by K. C. Chang. New Haven, Conn.: Yale University Press.
- Gillet, Philippe. 1993. *Le Goût et les mots: Littérature et gastronomie (XIV^e–XX^e siècles)*. Paris: Éditions Payot & Rivages.
- Girard, Alain. 1977 "Le Triomphe de La Cuisinière bourgeoise: Livres culinaires, cuisine et société en France aux XVII^e et XVIII^e siècles." *Revue d'Histoire Moderne et Contemporaine* 24 (October–December): 497–523.
- Goblot, Edmond. (1927) 1967. *La Barrière et le nouveau: Essai sur la bourgeoisie française moderne*. Paris: Presses Universitaires de France.
- Grimod de la Reynière, Alexandre Balthazar Laurent. 1806. *Journal des Gourmands et des Belles, ou L'Epicurien français*. Paris: Capelle-Renaud.
- . 1978. *Écrits gastronomiques*, edited by Jean-Claude Bonnet. Paris: Union générale d'Éditions.
- . (1808) 1983. *Manuel des Amphitryons: Contenant Un Traité de la Dissection des viandes à table, la Nomenclature des Menus les plus nouveaux pour chaque saison, et des Éléments de Politesse gourmande*. Paris: Éditions A. M. Métailié.
- . (1803–12) 1984. *Almanach des Gourmands servant de guide dans les moyens de faire excellente chère par un visé amateur*, 8 vols. Paris: Valmer.
- Guy, Christian. 1962. *Une histoire de la cuisine française*. Paris: Les Productions de Paris.
- Hobsbawm, Eric. 1983. "Inventing Traditions." "Mass-Producing Traditions: Europe, 1870–1914." Pp. 1–14, 263–307 in *The Invention of Tradition*, edited by Eric Hobsbawm and Terence Ranger. Cambridge: Cambridge University Press.
- Hyman, Philip, and Mary Hyman. 1997 "Imprimer la cuisine. Les livres de cuisine en France entre le XV^e et le XIX^e siècle." Pp. 643–55 in *Histoire de l'alimentation*, edited by Jean-Louis Flandrin and Massimo Montanari. Paris: Fayard.
- Larousse, Pierre. (1866–79) 1982. *Grand dictionnaire universel du XIX^e siècle*. Geneva-Paris: Slatkine Reprints.
- Mauss, Marcel. 1967. *The Gift*, translated by I. Cunnison. New York: Norton.
- Mennell, Stephen. 1985. *All Manners of Food: Eating and Taste in England and France from the Middle Ages to the Present*. Oxford: Basil Blackwell.
- Mercier, Louis-Sébastien. (1788) 1994. *Tableau de Paris*, 2 vols. Paris: Mercure de France.
- Mintz, Sidney W. 1996. *Tasting Food, Tasting Freedom: Excursions into Eating, Culture, and the Past*. Boston: Beacon Press.
- Morgan, Lady (Sydney). 1831. *France in 1829–30*, 2d ed. 2 vols. London: Saunders & Otley.
- Nora, Pierre. 1996–98. *Realms of Memory*, 3 vols. New York: Columbia University Press.
- Ory, Pascal. 1992. "La Gastronomie." Pp. 822–53 in *Les Lieux de mémoire: La Nation—Les France—II*, edited by Pierre Nora. Paris: Gallimard.
- Périgord, A. B. de (Horace Raison). 1825. *Nouvel Almanach des Gourmands servant de guide dans les moyens de faire excellente chère*. Paris: Baudouin Frères.
- Pitte, Jean-Robert. 1997. "Naissance et expansion des restaurants." Pp. 765–78 in *Histoire de l'alimentation*, edited by Jean-Louis Flandrin and Massimo Montanari. Paris: Fayard.

- Revel, Jean-François. 1979. *Un festin en paroles*. Paris: Pauvert.
- Ruoff, Marcel. (1924) 1994. *La Vie et la passion de Dodin-Bouffant-Gourmet*. Paris: Le Serpent à Plumes.
- Simmel, Georg. (1910) 1994. "The Sociology of the Meal," translated by M. Symons. *Food and Foodways* 5:345-50.
- Sue, Eugène. 1992. *Les Sept Péchés capitaux: La Gourmandise, Oeuvres complètes*, vol. 52. Geneva-Paris: Slatkine Reprints.
- Teuteberg, Hans-Jurgen, and Jean-Louis Flandrin. 1997. "Transformations de la consommation alimentaire." Pp. 725-46 in *Histoire de l'alimentation*, edited by Jean-Louis Flandrin and Massimo Montanari. Paris: Fayard.
- Viala, Alain. 1984. *Naissance de l'écrivain*. Paris: Minuit.
- Weber, Eugen. 1976. *From Peasants into Frenchmen*. Stanford, Calif.: Stanford University Press.
- Wenfu, Lu. 1988. *Vie et passion d'un gastronome chinois (Meishihji)*, translated by A. Curien and Feng Chen. Arles: Éditions Philippe Picquier-Unesco.
- Wheaton, Barbara. 1983. *Savoring the Past: The French Kitchen and Table from 1300 to 1789*. Philadelphia: University of Pennsylvania Press.
- Wolfe, Tom. 1975. *The Painted Word*. New York: Farrar, Straus & Giroux.
- Zukin, Sharon. 1991. *Landscapes of Power: From Detroit to Disney World*. Berkeley and Los Angeles: University of California Press.
- . 1995. *The Cultures of Cities*. Cambridge, Mass.: Blackwell.

Linking Action to Social Structure within a System: Social Capital within and between Subgroups¹

Kenneth A. Frank
Michigan State University

Jeffrey Y. Yasumoto
University of Chicago

Differences in transactions within and outside of cohesive subgroups are hypothesized to be a function of actors' pursuit of different forms of social capital. In an example of the French financial elite, subgroups are identified based on the pattern of friendships, and graphical representations establish the descriptive link between social structure and action. Estimates from multilevel models quantify the extent to which actors abstain from hostile action against subgroup members but tend to support others not in their subgroup. These complementary findings establish the subgroup as a critical mesolevel entity, defined by the social structure while affecting action.

One of the strongest and most enduring conceptualizations of the internal structure of systems is that of a set of integrated substructures. Classically, actors establish primary affiliations with members of their subunit, while uniquely defining their roles through ties with members of other subunits (Durkheim 1933; Granovetter 1973; Nadel 1957; Simmel 1955; Weber [1922] 1947). But accounts of how actors maintain ties within subunits to foster solidarity (e.g., Freeman 1992a; Hechter 1983; Merton 1957; Olmsted and Hare 1978), and between subunits to access diverse information or resources (Granovetter 1973; Selznick 1961; Simmel 1955), treat

¹ We thank Charles Bidwell, Nicole Ellefson, Kyle Fahrbach, Charles Kadushin, Steve Raudenbush, Becky Sandefur, Steve Sheldon, Ezra Zuckerman, and the *AJS* reviewers for their thoughtful comments on this work. This article is based on papers previously presented at the Social Networks Sunbelt convention and at a workshop session of the Social Capital Interest Group, at Michigan State University. We would like to thank those in attendance for the valuable feedback. Faults in the paper are our own. Please direct correspondence to Kenneth Frank, Room 460 Erickson Hall, Michigan State University, East Lansing, Michigan 48824-1034. E-mail: kenfrank@msu.edu

ties within and between subunits as from different realms; it is left to the individual to transition between realms to balance between the need for solidarity and the pursuit of diverse resources. Instead, we will argue that actors pursue a single resource, social capital, through different *mechanisms* according to the social structure in which any given action is embedded.

We will focus on one particular form of subunit, the cohesive subgroup, which defines a critical mesolevel in many systems of human actors. From the social-psychological perspective, individuals are most strongly influenced by members of their primary groups—people with whom they engage in frequent interactions (Cooley 1909; Epstein 1961; Festinger, Schachter, and Back 1950; Kadushin 1966)—and anthropologists have argued that primary groups are integral to understanding people within the contexts of their communities (Barnes 1972; Bott 1971). The corresponding sociological entity is the cohesive subgroup, with boundaries commonly defined for all subgroup members and for all actors in a system (Blau 1977; Homans 1950; Simmel 1955; Simon 1965; see Freeman [1992a] for a review).

At the level of the system, subgroups may define the essential components that contextualize actors' social ties and relations. For example, throughout this article we will draw on Lévi-Strauss's description of systems with a dual organization in which "members of a community . . . are divided into two parts which maintain complex relationships varying from open hostility to very close intimacy, and with which various forms of cooperation and rivalry are associated" (Lévi-Strauss 1969, p. 69). Lévi-Strauss's dual organization is comprised of kinship-based "moieties," which are comparable to "subgroups" in that ties are concentrated within the moieties. But like a subgroup, each moiety is also partially defined by reciprocated agreements and transactions with other moieties.² Thus, the boundaries of the moieties are not completely institutionalized but are permeable as sets of kinship ties becomes complex. Lévi-Strauss even allows for dual systems to have crosscutting or nested subpartitions.

Though Lévi-Strauss used examples from ancient and relatively isolated cultures, he characterized the dual organization as a general structural phenomenon, in which moieties have the capacity to regulate "relations among members of the community, forming a framework for the numerous mutual obligations between individuals and groups. They

² Lévi-Strauss (1969), e.g., described the moieties defined by the *morung* (men's houses representing subdivisions of a village) among the Konyak, the northern and eastern representatives of the Naga of Assam. These *morung* form a dual system through intermarriage and the corresponding reciprocation of gifts associated with marriage (p. 78).

strengthen the sense of social unity and at the same time encourage competition, thus stimulating the activities of the system" (paraphrased from Lévi-Strauss 1969, 78, who cited von Fürer-Haimendorf 1938). Given the capacity of the dual organization to strengthen a system by drawing on underlying social ties to reconcile competitive forces, it is not surprising that Lévi-Strauss's characterization is consistent with comparable characterizations of small systems (e.g., Durkheim 1933; Nadel 1957), large systems (e.g., Blau 1977), and formal organizations (e.g., Homans 1950; Simon 1965).

But like others (e.g., Festinger et al. 1950; Newcomb, Turner, and Converse 1965; Robinson 1981; Simon 1965) Lévi-Strauss's justification for the importance of cohesive subgroups is based on a functional explanation at the level of the system. This leaves unanswered how and why rational actors simultaneously sustain their subgroups and the linkages between them.

Several strands in the large body of work dedicated to the study of group solidarity inform our understanding of how and why actors sustain their subgroups. Groups are initially defined in terms of a concentration of ties among actors (Festinger et al. 1950; Freeman 1992a; Mudrack 1989). Group norms then evolve, most importantly including the norm of solidarity (Durkheim [1933] refers to this as "mechanical solidarity"). Actors who violate the norm of solidarity are then sanctioned or ostracized (Coleman 1990; Merton 1957). Therefore, rational actors typically adhere to the norm and remain committed to the group in order to reap the benefits of membership in the group (Bourdieu 1986; Hechter 1983; Merton 1957). Of course, the order of the formation of social ties and normative development is ambiguous (Homans 1950; Feld 1981). Balance theorists account for this ambiguity by arguing that norms and ties are consolidated in groups as individuals seek to interact with others who hold similar beliefs and who engage in interaction with common others (Cartwright and Harary 1956; Heider 1958; Newcomb 1961).

The importance of ties between subgroups was highlighted by Granovetter (1973). Granovetter argued that weak ties—those that are infrequent and extend beyond an actor's immediate social circle—provide access to unique influence and information that can give an actor important advantages in the pursuit of resources or position. Relatively recently, Granovetter's strength of weak ties has been refined. For example, Burt (1992) has drawn the implication that actors who bridge between subgroups have access to unique resources and information that makes them powerful brokers in a system, and Bian (1997) has emphasized the importance of ties that extend beyond the immediate social circle regardless of the frequency of contact.

Although there are explanations for why rational actors sustain group

solidarity through ties within their subgroups and diversify by engaging in ties with others not in their subgroups, there is no common framework for integrating how actors balance the demands of ties within and between subgroups or when they choose to activate one set of ties over another. In the next section, we hypothesize that actors enact subgroup solidarity while concurrently sustaining links between subgroups as they pursue a single commodity, social capital, through different mechanisms, depending on the social structure in which any given action is embedded. We then explore our hypotheses through an example, the system of friendships and actions among the French financial elite.

Following the development of our hypotheses, we employ two new methods to conduct our analyses. First, we use clustering and multidimensional scaling (MDS) techniques (Frank 1995, 1996) to identify and represent the social structure in terms of ties within and between subgroups. Second, we incorporate the subgroup structure into multilevel models of directed actions. In the discussion, we draw on the results obtained through these methods to characterize the subgroup as a mesolevel entity that mediates between dyadic-level action and systemic coherence.

SOCIAL CAPITAL: THE COMMON COMMODITY

Social capital is defined in terms of resources that actors may access through social ties (Bourdieu 1986; Coleman 1988, 1990; Putnam 1993; Schlicht 1984; see Portes [1998] or Woolcock [1998] for reviews of the definition and development of the concept of social capital). These ties may affect one actor's action directed toward another based on the social structure in which the action is embedded (cf. Granovetter 1985) and the history of transactions between the actors. For example, the social structure may make one actor feel obligated to reciprocate a favorable action of another. The actor to whom another is obligated then possesses social capital.

A key aspect of social capital is that it is an "aid toward making micro-macro transitions" (Coleman 1988, p. S101). The behavior of individuals may be explained partly by their efforts to accumulate and draw on social capital, while the coherence of systems is sustained by the flow of resources and favors facilitated by social capital. But analyses based on theories of social capital have employed only the crudest transitions from micro to macro. Case studies of whole systems have typically transitioned from micro to macro by establishing evidence of social capital in the aggregate of actors' actions, ignoring internal variation in social structure and internal variation in actions embedded in the social structure (Coleman 1988; Granovetter 1985; Portes and Sensenbrenner 1993). On the other hand, statistical inferences based on models, including social capital at the level

of the individual, employ assumptions that actors are independent (e.g., Bian 1997; Uzzi 1996), ignoring the interwoven relations among actors that constitute systemic structure. The limitation is contained within Coleman's initial conception, as the above quote is completed: "without elaborating the social structure through which this occurs" (Coleman 1988, p. S101). In contrast, we believe it is essential to elaborate the social structure, defined by subgroup boundaries, in order to differentiate the contexts in which the pursuit of social capital is embedded.

We emphasize the distinction between two mechanisms through which actors pursue social capital, "reciprocity transactions" and "enforceable trust," which are sustained by different norms and patterns of social ties among actors (Portes and Sensenbrenner 1993). First, actors may pursue social capital directly through reciprocity transactions, which Portes and Sensenbrenner (1993), building on Simmel (1955), described as the "accumulation of 'chits,' " which actors hope will be reciprocated (p. 1324; see also Coleman 1988; Dawes 1988). Note that reciprocation is encouraged by a norm of reciprocity, which is sustained by at least moderately dense social ties (Portes and Sensenbrenner 1993). Of course, norms and behaviors coevolve as the norm of reciprocity is sustained by actors hoping to create an obligation in others whom they have supported.

But actors also may pursue social capital through substantive rationality involving particularistic obligations (Weber 1947), which was modified and renamed by Portes and Sensenbrenner (1993) as enforceable trust, the mechanism through which "social capital is generated by individual members' disciplined compliance with group expectations" (p. 1325). The language of enforceable trust suggests a constraint imposed upon an actor (Coleman 1990). That is, an actor refrains from acting against another who enforces her trust, foregoing a potentially immediate gain in resources (Uzzi 1996). Unlike the actor who anticipates that a direct supportive act will be reciprocated because of a group norm, the actor who enforces the trust of another through mediated social structure is assured only that the other will not act against her.

Actors are more likely to pursue social capital through enforceable trust when the social structure is dense, which facilitates awareness of implicit investments of social capital through abstention from hostile action (see Raub and Weesie [1990] on the reputation of actors embedded within densely woven networks). Further, actors establish a norm of solidarity by sanctioning acts of betrayal through the dense social structure (e.g., Coleman 1990; Festinger et al. 1950; Homans 1950; Portes and Sensenbrenner 1993). Therefore, "the actor's behavior is not oriented to a particular other but to the web of social networks" (Portes and Sensenbrenner 1993, p. 1325). And in the specific system of the French financial elite, Kadushin (1995) initially argued that enforceable trust "cannot be an attri-

bute of friendship pairs, nor is it deducible from the possession of common social attributes. Rather it stems from an *interwoven network* in which there is a clear expectation that actions will have positive and negative sanctions—not necessarily from one's immediate friends, but from a more diffuse friendship circle" (p. 219; emphasis added).

Portes and Sensenbrenner have done more than merely differentiate the classical mechanism through which actors establish trust associated with reciprocity transactions from trust that is *enforced* within subgroups. They have linked the type of trust to the social structure in which action is embedded and thus have consolidated both mechanisms under the rubric of social capital. But while Portes and Sensenbrenner have addressed differences in the mechanisms through which actors pursue social capital in different immigrant populations, we believe their theoretical distinctions are acute enough to differentiate the mechanisms through which actors are likely to pursue social capital *within* systems. Thus, though studies of social capital are commonly employed at a national or international level (see Woolcock [1998], esp. pp. 193–96, n. 20), we employ social capital here to the study of small systems so that we may more directly address the specific phenomena through which social capital is generated. In this sense we turn the theories of Portes and Sensenbrenner outside in.

We note that, in the internal social structure of small systems as delineated by the boundaries of cohesive subgroups, ties are sufficiently dense to sustain enforceable trust, while ties between subgroups may not be, which leads to the following hypothesis:

HYPOTHESIS 1.—*Actors are more likely to enforce each others' trust if they are members of the same subgroup than if they are not members of the same subgroup.*

Between subgroups, the ties may be too sparse to sustain enforceable trust. In the absence of enforceable trust, actors may pursue social capital primarily through reciprocity transactions. Put differently, the potential for enforceable trust within subgroups reduces the relative need to pursue social capital through reciprocity transactions within subgroups. This leads to our second hypothesis:

HYPOTHESIS 2.—*Ceteris paribus, actors are more likely to engage in reciprocity transactions if they are not members of the same subgroup than if they are members of the same subgroup.*

Direct ties and other foci may well support reciprocity transactions within subgroups, but the residual grouping effect is to decrease the relative tendency for reciprocity transactions within subgroups.

We have hypothesized a distinction according to subgroup boundaries; actors rely relatively more on enforceable trust within subgroups and reciprocity transactions between subgroups. This leads to a third hypothesis that synthesizes the first two:

HYPOTHESIS 3.—Actors are more likely to enforce each other's trust, versus engage in reciprocity transactions, if they are members of the same subgroup than if they are not members of the same subgroup.

Thus, we hypothesize that the subgroup boundary, originally defined in terms of the social structure, generally affects the types of actions in which actors engage.

In the next section, we introduce the system of the French financial elite, arguing that the social structure has an effect even on the actions of actors who are known for their aggressive pursuit of financial capital (De Quillacq 1992). We describe the exogenous and historical conditions in which much of the social structure was formed and present an overarching anecdotal link between the social structure and transactions among the actors. This sets the stage for a more complete analysis of the link between social structure and actors' actions.

THE FRENCH FINANCIAL ELITE

Background

The study of the system defined by the French financial elite and the relations and actions among them is valuable on several levels.³ First, the case study of the French financial elite is of interest in its own right. Not only does the system reflect the influences of the French government ministries and educational and political institutions, but members of the French financial elite enjoy a certain celebrity based on the prestige they acquired through their training and service (De Quillacq 1992; Kadushin 1995). Second, like other systems of elites, they exert considerable influence over their society's politics and economy, and so the internal machinations of this system shape the French economic and political landscapes (e.g., DiMaggio 1992; Higley et al. 1991; Laumann and Pappi 1976). Third, regardless of their connection to established French institutions, members of the French financial elite are typical of many actors in other systems defined by a competitive environment (e.g., the diamond traders described by Coleman [1990] and Granovetter [1985]).⁴

Kadushin (1995) operationally defined the French financial elite as the

³ We define the French financial elite as a system because the members are interrelated through their ties and actions but exhibit neither the formal structures and common goals of an organization, nor are goods frequently exchanged among them as in a market.

⁴ The study of the French financial elite is also opportunistic. Because of their celebrity, much is known about them regarding background, friendships, and actions. This information is critical to our analysis and is difficult to obtain for actors in a nonexperimental setting.

28 (all male) actors with highest aggregate prominence (Burt 1982), in terms of influence and recognition, from a larger snowball sample. The initial sample consisted of 125 people who held positions in major financial institutions and who were in press reports in the private sector (the data were originally presented in *The Power Brokers* [De Quillacq 1992]; see Kadushin 1995, pp. 206–7). More substantively, members of the French financial elite obtain their status by attending one of the *grand écoles*, in particular, the École National D'Administration (ENA), the most prestigious school for preparing people for “administrative” positions. Traditionally, the most successful students at ENA are selected to serve in high-level positions in what are called the *grand corps* of the French civil service, most of which are related to the finance ministry. Ultimately, the top individuals from the *grand corps* obtain high-level positions in the French Treasury. From there, they move to positions as CEOs in the public or private sector. Thus, unlike the immigrant populations described by Coleman (1990), Granovetter (1985), and Portes and Sensenbrenner (1993), there is a continuity to the experiences of the French financial elite from their education, to their early employment, and then to their later careers.

Because of the government's control over employment opportunities and the distribution of resources to financial institutions, the political affiliations of members of the French financial elite affect their success in the system at any given time (De Quillacq 1992; Kadushin 1995). In turn, members of the French financial elite who are affiliated with a common political party may develop social capital through “bounded solidarity,” which refers to actors of a “particular group who find themselves affected by common events in a particular time and place” (Portes and Sensenbrenner 1993, p. 1327). More generally, we believe that the openness of the French financial elite to a vacillating political system requires the actors to diversify their investments of social capital.

The powerful educational and government institutions not only serve to select the French financial elite and influence their success, but they also cultivate the social structure among them. Many of the members of the French financial elite form lasting friendships while attending ENA, which partly stresses the development of relationships with one's peers (De Quillacq 1992, chap. 2). These patterns of friendship then consolidate with political allegiances as the actors trade on their educational backgrounds to obtain high-level posts in government ministries (De Quillacq 1992, chap. 3; Kadushin 1995).

The social structure experienced by the members of the French financial elite represents an inversion of the development of Simmel's metropolitan man. As Simmel described it, the metropolitan man preferred the overlapping social circles experienced in the city setting, where an individual

could uniquely define himself through membership in multiple groups and organizations that did not themselves overlap. In contrast, educational and employment backgrounds are consolidated in the social circles of the French financial elite, constructing a small world for these most metropolitan of men. As one member of the French financial elite stated: "Dinners, Luncheons, breakfasts, tête-à-tête. . . . It's always the same ones who talk, always the same ones who are there. It doesn't stop. We meet all the time" (Kadushin 1995, p. 210).

Not surprising, the actions of members of the French financial elite in their positions as CEOs are linked to the friendship structure established early in their careers. For example, one actor is reported to have supported another "mainly out of friendship" (De Quillacq 1992, p. 103).⁵ In this sense, the system of the French financial elite is similar to other systems of elites (e.g., Laumann and Pappi 1976).

But action can be linked to a more general social structure. For example, as one venerable French banker stated: "When it comes to doing business in France the basic principle is to pass each other the 'rhubarb.' I give you some business, you give me some. . . . If you want to do business with me, you get hold of one of my closest friends. Since the market doesn't count—because the state regulates the market place—it's friendships that are important" (De Quillacq 1992, p. 67).

The banker's comments reveal that transactions are embedded in a very small context of social relations. The first statement—"I give you some business, you give me some"—is consistent with the definition of reciprocity transactions. But the transactions are also linked to social contexts of a few actors, as indicated by another phrase: "If you want to do business with me, you get hold of one of my closest friends." Finally, the banker confirms the extent to which action is embedded in the overall social context of the French financial elite—"it's friendships that are important"—and links this to the general political context in which the French financial elite operate—"the market doesn't count—because the state regulates the market place."⁶ Part of our goal is to refine this link between action and social structure. In particular, we will follow Kadushin (1995) in characterizing the system based on Lévi-Strauss's (1969)

⁵ This alignment of action and friendship represents the accumulation of social capital through a basic mechanism, "value introjection," which "prompts individuals to behave in ways other than naked greed" (Portes and Sensenbrenner 1993, p. 1323).

⁶ The importance of friendships amidst a state-dominated economic sector is confirmed by several articles in the *Economist* (*Economist* 1992b, 1992d, 1992e, 1993b, 1993c, 1993d, 1994a, 1994b, 1995c, 1995d). This pattern also is found in Bian's (1997) connection between the importance of social ties and employment in China in the absence of a market mechanism.

dual organization to help us understand how the inherent competition among the French financial elite is embedded in their social ties.

The Social Structure of the French Financial Elite

We use Frank's (1995) *KliqueFinder* algorithm to identify cohesive subgroups in order to characterize the pattern of friendship ties among the French financial elite.⁷ Frank's algorithm employs a stochastic definition of cohesion, emphasizing ties within subgroups *relative to* actors' propensities for engaging in ties and sizes of subgroups.⁸

Typically, the criteria used to define cohesiveness are fixed (i.e., in terms of the absolute number of ties present or absent in a subgroup), and the corresponding structures accommodate variation in the data by allowing the subgroup boundaries to overlap, producing what Kadushin (1995) referred to in his original analysis of friendships among the French financial elite as "several overlapping small clusters" (p. 212; see also Freeman 1992a; Frank 1993). Such overlapping clusters are inconsistent with many of the theoretical characterizations of systems composed of nonoverlapping subgroups. In particular, overlapping boundaries fail to establish "an inside and an outside" (Abbott 1996, p. 872) necessary to define a sociological entity. How can we differentiate action within subgroups from that between subgroups if two actors can simultaneously be members of the same subgroup and members of different subgroups? At the more macrolevel, though ties between members of different subgroups can and must exist, systemic processes, such as those associated with Lévi-Strauss's (1969) dual organization, cannot be characterized in terms of components if the components are not distinct.

In table 1, we show the nonoverlapping subgroups identified by Frank's algorithm. The elements in the cells of the matrix in table 1 take a value of 1 if the members of the French financial elite in the corresponding

⁷ Measures of friendships were obtained through interviews conducted by the French journalist, De Quillacq. Friendships were treated, for the most part, symmetrically—if *i* listed *j* as a friend, then it was assumed that the relationship was reciprocal. See Kadushin (1995, 210 n. 10) for details.

⁸ Frank's criterion can be defined as the increase in the odds that there will be a tie between two actors (as opposed to the absence of a tie) that is associated with membership in a common subgroup. The odds ratio is large to the extent that actors engage in ties within subgroups and do not engage in ties between subgroups. The ratio is small to the extent that actors do not engage in ties within subgroups and engage in ties between subgroups. Balancing the two effectively accounts for subgroup sizes and the propensities for actors to engage in ties. Thus, maximization of the criterion described by Frank satisfies Kadushin's (1995) call for an algorithm that "maximizes both in-[sub]group density and minimizes out-[sub]group relations" (p. 212).

TABLE 1
PARTITIONED FRIENDSHIPS AMONG THE FRENCH FINANCIAL ELITE

Actor	GROUP																												
	A							B							C							D							
	8	9	12	19	20	21	24	25	26	27	2	5	6	10	13	14	17	1	4	15	16	18	22	28	3	7	11	23	
Group A:																													
8	1	.	1	1	.	.	.	1	.	.	1	1	.	.	.
9	1	.	.	1	1	1	1
12	1	1	1	1	1	1
19	1	1
20	1	1
21	.	1	1	1	.	.	1
24	1	.	1	1	1	1	1
25	.	.	1	.	.	.	1	.	1	1	1
26	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
27	1	1	1	1	.	1	1	1	1	1	1	1	.	1	.	.	.
Group B																													
2	.	1	1	.	.	1	.	.	.	1	.	1	.	1
5	1	1	1	1

6	1	1
10	1	1	1
13	1
14	1	1	1	1	1
17	1	1	1	1	1	1	.	.	1
Group C:																			
1	1	.	1	1	1	.
4	.	.	1	1	.	.	.	1	1	1	.	1	.
15	1
16	1
18	1	.	.	.	1
22	1	1	.	.	.
28	1	1	1	.	.	.
Group D:																			
3	1	.	1	1	1	1	.	.	1	1
7	1	1	.
11	1	.
23	1	.

NOTE— $N = 28$.

row and column indicated being friends.⁹ For example, actors 8 and 24 of subgroup *A* reported being friends. Four subgroups were identified, with 10 actors in subgroup *A* (8, 9, 12, 19, 20, 21, 24, 25, 26, 27), seven in subgroup *B* (2, 5, 6, 10, 13, 14, 17), seven in subgroup *C* (1, 4, 15, 16, 18, 22, 28), and four in subgroup *D* (3, 7, 11, 23). One can observe a concentration of friendships along the block diagonal of the matrix in table 1 (the odds that two actors are friends is more than 10 times greater within subgroups than between subgroups), which is consistent with the characterization of a system composed of cohesive subgroups.

We use Frank's (1995) simulations to address two specific questions regarding the identified subgroup memberships. First, because application of the algorithm would reveal a concentration of ties within subgroups in data in which the ties were formed at random, we must ask whether the friendships are more concentrated within the subgroup boundaries in table 1 than is likely to have occurred by chance alone.¹⁰ In this case, the monte carlo generated probability of obtaining the concentration of friendships within the subgroup boundaries in table 1, *after employing the clustering algorithm*, was less than one in 1,000 by chance alone ($P \leq 0.001$). That is, although, as Kadushin noted, the density of friendships was high among the French financial elite as a whole, there is strong evidence that friendships were concentrated within the boundaries of cohesive subgroups.

Second, we ask whether the algorithm accurately recovered the "true" subgroup memberships.¹¹ For networks of comparable size and concentration of ties within subgroups as in table 1, Frank's simulations indicated that the odds that two actors in the same known subgroup were assigned by the algorithm to the same observed subgroup were five times greater than they were for two actors in different known subgroups. This finding, taken with the evidence that ties are concentrated within subgroups, supports a conclusion that the subgroup boundaries in table 1 have not been imposed on a fluid pattern of friendships but represent an empirical ten-

⁹ The identification numbers for the actors were assigned according to their rank of influence as reported by the other actors. Actor 1 was reported by others to have the most influence, actor 2 the second most, and so on. Only the ranks for the top 20 actors were known, so the last eight, numbers 21 through 28, were assigned arbitrarily.

¹⁰ In these simulations, we record the odds ratios after repeated applications of the clustering algorithm to networks of the same size and distribution of row marginals (the out-degrees are considered fixed), as in table 1, but in which the ties among the actors were randomly assigned.

¹¹ To allow others to address this question, Frank (1995) applied his algorithm to multiple simulated networks in which the subgroup memberships were known. Frank then compared the subgroup memberships identified by the algorithm with the known subgroup memberships

dency for the friendships among the French financial elite to be concentrated within the subgroup boundaries displayed in table 1.

To more fully render the structure of friendships in table 1, we used the subgroup boundaries to locate the actors as shown in the *crystallized* sociogram in figure 1 (see Frank [1996] for the technique of applying MDS within and between subgroups to generate an image such as fig. 1). We refer to the sociogram as *crystallized* because the substructures are defined by high concentrations of ties and then integrated through sparse ties. The technique reorganizes the data to generate a more accessible, clearer image of the social structure that is consistent with characterizations of systemic structure such as Lévi-Strauss's dual organization.¹²

Figure 1 reveals the value of using a stochastic criterion to define cohesion. Instead of identifying several overlapping subgroup boundaries according to a fixed criterion, we identified a few nonoverlapping but *permeable* subgroup boundaries (consistent, e.g., with Lévi-Strauss's description of moieties). The subgroup boundaries then delineate the four components (A, B, C, and D) of the internal structure of the system as well as define the contexts for actors' actions.

To characterize the subgroups in the system of the French financial elite, we begin by identifying the focus (Feld 1981) of each subgroup in terms of conditions external to the system. (Note that we recognize Feld's argument that the chronological sequence of the formation of ties and the emergence of a focus is ambiguous.) Following Kadushin (1995) and Frank and Yasumoto (1996), we include information in our sociogram

¹² Admittedly, the image in fig. 1 imposes a distortion on the pattern of friendships by locating actors most closely with members of their cohesive subgroups (although in this case the ties are so concentrated within subgroups that we had to expand the area within the subgroup boundaries by a factor of six to adequately represent within subgroup patterns of ties in the same image as the sparse ties between subgroups). In general, this is an acceptable distortion to the extent that subgroup boundaries capture unusually dense regions of ties. And yet, the distortion imposed may be too extreme in cases where the odds ratio is only slightly affected by the assignment of an actor to one subgroup over another. This could be addressed by unaffiliating actors who fit almost equally well in more than one subgroup. In the present case, we have unaffiliated actors who have only one friendship with a member of their subgroup and a friend in at least one other subgroup. Such actors are clearly not strongly affiliated with any subgroup in the system, since their initial assignment was established through a single friendship. Therefore, actors 1, 16, and 28, originally of subgroup C, and actor 7, originally of subgroup D, appear unaffiliated in fig. 1. We do not make a strong claim for use of this criterion of "disaffiliation." We note only that the sparseness of ties among the original members of subgroup C was inconsistent with the conditions necessary to sustain enforceable trust. Our interpretations would not be affected were another criterion to be used provided actor 1 is ultimately not affiliated with the members of subgroup C (actor 1 is a friend of only one other in subgroup C, and that one other has only one friend in subgroup C).

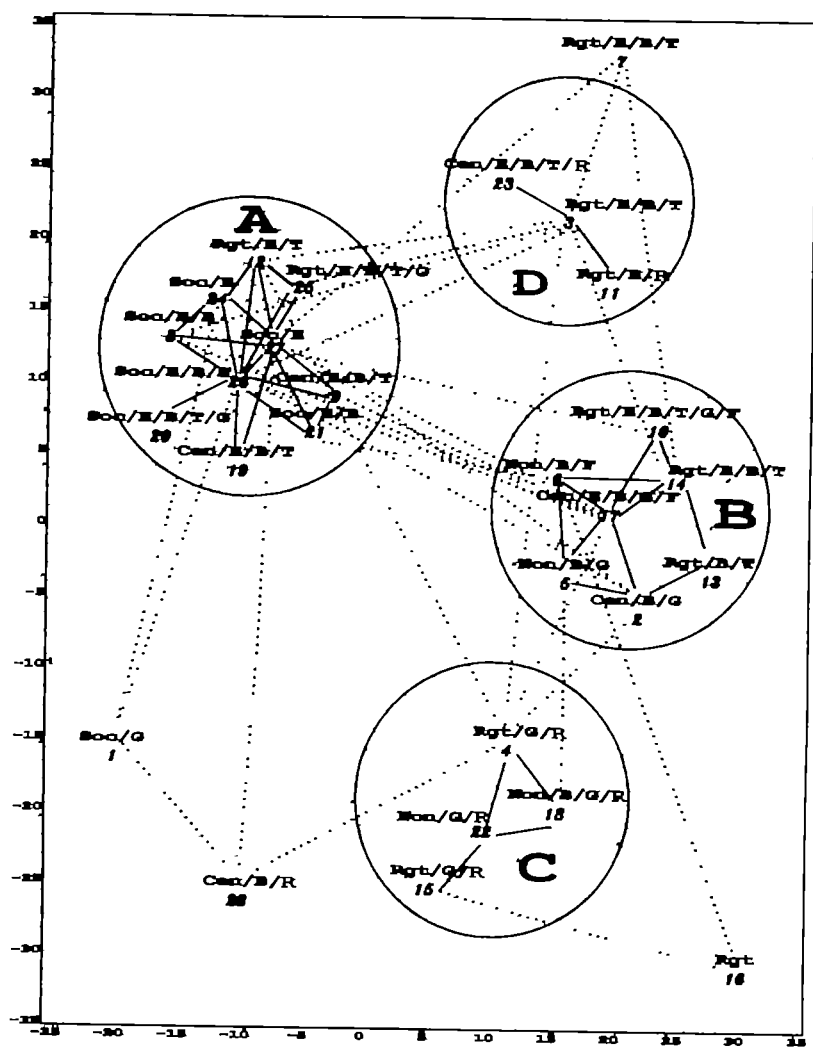


FIG. 1.—Crystallized sociogram: friendships among the French financial elite (solid lines indicate ties within subgroups and dotted lines, ties between subgroups; scale = $1/[\text{density of ties}]$; within subgroups, scale = $6/[\text{density of ties}]$).

regarding each actor's party affiliation (Soc = Socialist, Cen = Central, Rgt = Right, and Non = None, or unknown), as well as hostility to the Socialists, using an "R" to indicate membership in *Resistance de Socialism*. We have also represented whether or not the actor attended the prestigious *Ecole National d'Administration* (ENA) as indicated by an "E." Because the context for action also depends on the financial institutions with which the actors have been affiliated, we indicate whether the actor was ever employed in the treasury as indicated by a "T," whether the actor was ever employed in the banking industry as indicated by a "B," whether the actor was ever employed in *Grand Banque*, one of the larger banks in France, by a "G," and whether the actor was employed in the prestigious *Firme de la Finance* by an "F."¹³ (This information was obtained directly from Charles Kadushin and from De Quillacq [1992].)

In figure 1, each subgroup's focus may be linked to external political affiliations and restricted access to educational institutions, particularly ENA. There is a core set of Socialists who attended the elite ENA (subgroup A), an opposition with allegiance to the political Right, many of whom also attended ENA (subgroup D), a politically neutral subgroup of bankers (subgroup B), a set of old guard bankers hostile to the Socialists (subgroup C), and a set of unaffiliated actors, 1, 7, 16, and 28, who, except for actor 7, did not have the opportunity to establish a host of friendships through attendance at ENA or through service in a formal government ministry. If the system were not as reflective of cultural influences, then the system may not contain subgroups A and D whose members established friendships during attendance at ENA (Bourdieu 1986). If the system were not open to the influence of political institutions, we may not have identified subgroup A, which contains the Socialists, as well as subgroup C, whose members share an animosity toward the Socialists.

Our representation in figure 1 is consistent with Kadushin's (1995) characterization of the system as a dual organization. In particular, Kadushin's left moiety (as defined in Kadushin 1995 p. 211, fig. 1) appears at the top of our figure, and his right moiety appears at the bottom.¹⁴ Therefore, it

¹³ "Resistance de Socialism," "Grande Banque," and "Firme de la Finance" are terms used to disguise the identity of the individual actors. We did not use pseudonyms for larger organizations, and we occasionally reveal specific details that are relevant to our hypotheses. We believe we have more latitude to reveal information than did Kadushin partly because more time has elapsed since the data were obtained.

¹⁴ Kadushin first attempted to identify cohesive subgroups, or cliques, in the friendship pattern. Upon the failure of the various cohesion-based algorithms to identify nonoverlapping subgroups, Kadushin identified the halves of the moiety based on consensual results of several algorithms employing structural-equivalence-based criteria. Our use of a stochastic criterion has allowed us to identify cohesive subgroups consistent with Kadushin's initial instinct and with the characterization of the system as a dual organization.

is not surprising that we confirm Kadushin's findings that party affiliation and graduation from ENA were related to block membership. But our image delineates regions of dense and sparse ties more explicitly than Kadushin's figure 1. This allows us to more directly link the mechanisms through which individuals pursue social capital to the social structure in which their action is embedded.¹⁵

In the next section, we offer examples in which actors' actions reflect and generate the norms needed to sustain the accumulation of social capital. In graphical displays and multilevel models, we then establish the link between the social structure, as represented by subgroup membership, and hostile and supportive directed actions. It is our intention to demonstrate that, taken as a set, our findings are consistent with the hypothesis that actors pursue a single commodity, social capital, through different mechanisms according to the social structure.

THE LINK BETWEEN THE SUBGROUP STRUCTURE AND DIRECTED ACTIONS

Throughout the 1980s and 1990s, the French economic and political systems fluctuated between economic nationalization, as advocated by Socialists, and privatization, as advocated by members of the political Right (recall Mitterand's 1988 presidential election campaign based on the "ni-ni" policy, which called for neither more privatization nor more nationalization in the hopes of stabilizing the economy). These fluctuations generated discontinuities in the economic system, such as changes in policies regarding the percentage that the state could hold in any company. Dominating the board memberships and CEO positions of the French financial ministries and the major French banks and insurance companies as of the early 1990s, members of the French financial elite formed many short-term alliances and launched quick strikes as they seized opportunities created by these discontinuities. These alliances and attacks define the actions that we link to social structure in the following sections.

Hostile Action and Enforceable Trust

We begin by exploring the link between subgroup membership and the extent to which actors pursue social capital through enforceable trust. To

¹⁵ Our data will not allow us to easily infer the intent of the actor or the processes that encompass their actions, both of which are needed to substantiate our use of the term "pursue." We use the term to be consistent with our theory and because it reflects the fact that the goal of accumulating social capital may not be achieved. In order to be more accurate, we might write, "action consistent with the pursuit of social capital," but we find this an awkward phrase.

establish that the conditions necessary for enforceable trust to arise were satisfied, we note instances in which an act of hostility elicited negative sanctions, thus reflecting and perpetuating a norm of solidarity (our emphasis on negative sanctions is consistent with Coleman 1990). Data regarding actions of the French financial elite toward one another (which we will call "directed action") from 1980 to 1990 were obtained from reports in De Quillacq (1992), and data regarding actions from 1990 to 1995 were obtained from the *Economist* (1990–95).¹⁶ An act was characterized as hostile if its execution would have denied another actor access to resources or formal position. As examples of hostile action, actor 3 attempted to take over a large bank headed by actor 5, and actor 18 helped oust actor 9 from his position as the head of Grande Banque.¹⁷

Of the 15 hostile acts identified, although some appear to be part of ongoing personal conflicts, most did not elicit negative sanctions from those not directly involved in the conflict. Two examples, however, are instructive. First, Alain Minc is now known as a "red flag in the face of the establishment" (De Quillacq 1992, p. 88), partly because he helped engineer a bid for a large private conglomerate that was held by members of the financial elite. Ever since this failed bid, Minc has "probably been a marked man in the eyes of the establishment" (De Quillacq 1992, p. 89) and was not included in the list of most influential actors. He has been ostracized, limiting his status and access to resources.

We take this example of a norm of solidarity along with the overall dearth of hostile acts over the 15-year period (1980–95) as partial support for Kadushin's claim that there is at least a moderate level of enforceable trust throughout the system. Although there were almost monthly oppor-

¹⁶ The first author coded all the data.

¹⁷ Our record of actions may not be exhaustive. In this sense, we treat the reported actions as a sample from all actions over the duration of the system. Given that our data were initially reported by journalists, the events of which we are aware are most likely more "newsworthy" than the average event in the system. To this extent, we expect that any bias would run counter to our first hypothesis, in favor of a failure of enforceable trust within subgroups, since such failures represent greater transgressions and are therefore more newsworthy. Because the subgroup boundaries represented in fig. 1 were not known to journalists in general, it is difficult to support our claim directly. But we note that the *Economist* has devoted several articles to strife generally within the higher levels of political parties and government cabinets (*Economist* 1991, 1992a, 1992c, 1992d, 1993a, 1995b). Apparently, journalists are able to find, and will report, evidence of strife within fairly closed political circles. We expect that journalists' abilities to report strife within political circles would apply to the subgroups among the French financial elite, which are themselves focused partly around political affiliations. Further, the interest in reporting such strife for the French financial elite would be comparable to the interest in reporting on internal strife among politicians because of the renown of the French financial elite (Kadushin 1995).

tunities for these actors to undermine one another, that we know of only 15 such acts suggests that members of the French financial elite have curtailed their potentially hostile actions because of fear of negative sanctions. Indeed, the French financial elite are often described in De Quillacq (1992) as a "old boy" network, partly based on their reserve in dealing with one another.

The second example of the administration of negative sanctions in response to a hostile act reveals the norm of solidarity of subgroup members. Actor 27 (among the Socialist graduates of ENA in subgroup *A*), who was head of a partly nationalized bank, asked actor 1, who was head of a large insurance firm that had considerable financial resources, to help in a bid for a private financial firm. Actor 1 refused, even after pressure from the prime minister, and effectively blocked the action of 27. The bid failed miserably and publicly, costing actor 27 standing in the system. The friendship between actors 1 and 27 was strained during the late 1980s, and the two did not list each other as friends in the data used to construct table 1. Most important, prior to this episode, most of actor 1's ties were with members of subgroup *A*, but now actor 1 is described as a "traitor to socialism" (De Quillacq 1992, p. 131). Actor 1's decision to act independently has cost him membership in a cohesive subgroup, establishing a precedent for those who consider violating the trust of subgroup members.¹⁸

Although the examples of Alan Minc and actor 1 reveal the consequences of the failure of enforceable trust, they reveal sanctions imposed to establish a norm of solidarity. This norm then applies to the actions of others in the system. For example, actor 17, a very popular actor, is located in the center of subgroup *B*. This would seem to make him privy to unique resources and information on which he could trade to become a very influential actor (See Freeman 1979). And yet, he is likely constrained by the norm of solidarity applied to and through the friendships in which he is enmeshed—his trust is enforced by the other actors in the system in general and by his subgroup members in particular. Thus, actor 17 has been described as being "too nice to be number one" (De Quillacq 1992, p. 380).

The concentration of ties within subgroups can sustain a stronger norm of solidarity than across the system as a whole. The sparse ties between subgroups may well account for why there were few sanctions for most of the hostile actions, most of which were taken against unaffiliated actors or against a member of another subgroup. In fact, the attitude of members

¹⁸ On the other hand, actor 1 is characterized as one of the most influential actors in the system. In a subsequent section, we will show how he has achieved this status partly through reciprocity transactions.

of the system was typically to condone hostile action crossing subgroup boundaries as inherent to a competitive system. For example, note the lack of moral or ethical concern in one actor's discussion of a hostile action directed by one actor toward another of a different subgroup: "*Most of all I found that it was very badly executed. . . . The affair was very poorly mounted from a technical point of view*" (De Quillacq 1992, p. 105; emphasis added).

To represent the link between action and the old boy networks, in figure 2 we display the 15 reported hostile acts within the geometric space defined by figure 1.¹⁹ Critically, the *only* example of a hostile act taken by one actor against a subgroup member occurred when actor 10 of subgroup B resisted the entrance of actor 6's son-in-law into *Firme de la Finance*. Even this example is mild relative to other hostile actions.

The trend observed in the graphic is consistent with a rephrasing of hypothesis 1 in terms of the relationship between hostile action and subgroup membership among the French financial elite:

RESTATEMENT OF HYPOTHESIS 1.—*Among the French financial elite, actors are more likely to abstain from hostile action (abstinence from hostile action being consistent with enforced trust) if they are members of the same subgroup than if they are not members of the same subgroup.*

Note that in our data we are not able to directly address whether actors pursue social capital. To do so would require extensive data regarding actors' states of mind, expectations, and longitudinal data with regard to choices to engage in and abstain from hostile action. Such data are beyond the scope of what was available for these actors. Thus, we rely on the pattern of hostile actions and indications of general norms of solidarity as rough indicators of enforceable trust.

We will test our hypothesis by estimating the effect of subgroup membership on the logit of the probability of a hostile act being initiated by i toward i' as in the following model:

$$\log \left(\frac{P[HA_{ii'} = 1]}{1 - P[HA_{ii'} = 1]} \right) = \beta_0 + \beta_1 (MMS_{ii'}), \quad (1)$$

¹⁹ The key for the hostile actions represented in fig. 2 is: $i \rightarrow i'$ indicates that actor i aggressively displaced actor i' , such as when actor 18 helped to oust actor 9 from *Grande Banque* and then took over at *Grande Banque*. $i \uparrow i'$ indicates that i investigated i' , such as when actor 4 investigated actor 10 and his agency. $|i| \leftarrow i'$ indicates that actor i impeded some action of actor i' , such as when actor 13 helped to prevent actor 3's attempted takeover of a large bank (headed by actor 5). $|i| \leftarrow i' (P)$ indicates that actor i impeded a policy of actor i' , such as actor 27 inhibiting actor 1's attempts to distribute the funds of his bank for various social policies.

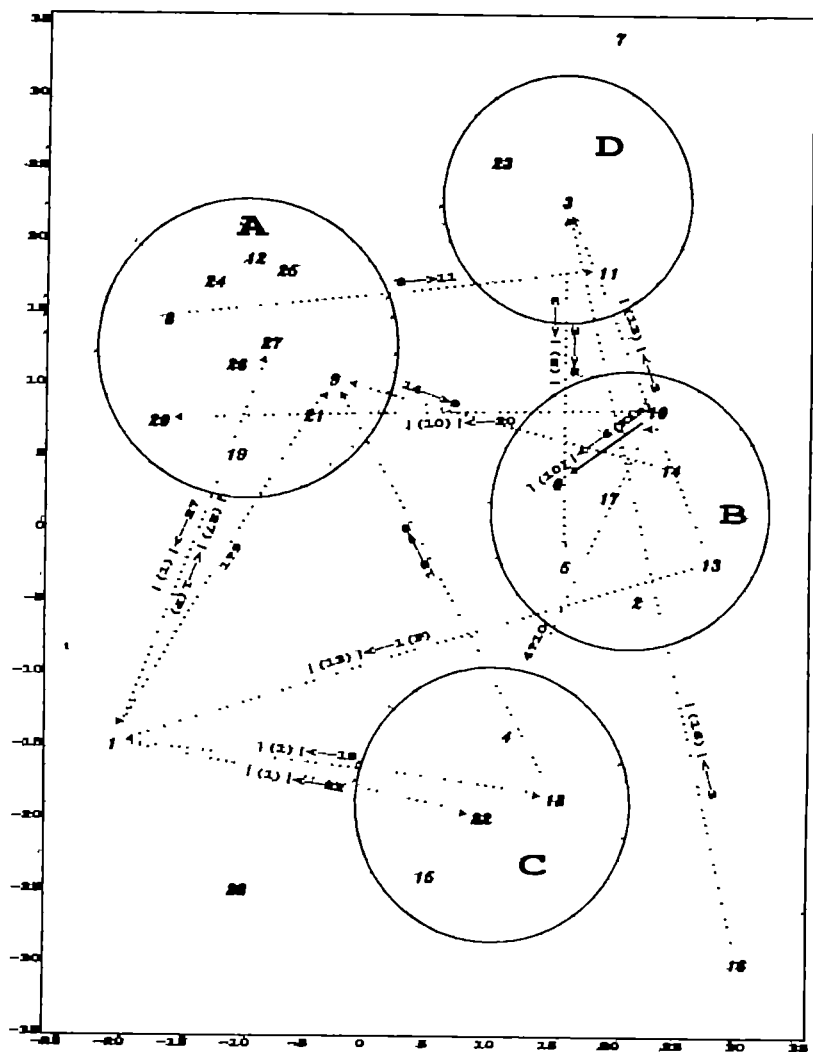


FIG. 2.—Hostile acts embedded in the friendship structure (solid lines indicate acts within subgroups, and dotted lines, acts between subgroups).

where:

$$\begin{aligned}
 HA_{i,i'} &= 1 && \text{if actor } i \text{ engaged in a hostile act towards } i', \\
 &= 0 && \text{otherwise,} \\
 MMS_{i,i'} &= 1 && \text{if actors } i \text{ and } i' \text{ are members of the same subgroup,} \\
 &= 0 && \text{otherwise.}
 \end{aligned}$$

As we model the link between social structure and directed actions, we extend the work of those who have linked social structure to actors' attributes at the individual level (e.g., Coleman 1988; Frank 1996; Kadushin 1995; Uzzi 1996), including analyses based on systemic level theories (e.g., Burt 1992). Further, we extend models of social and economic support among dyads (e.g., Wellman and Wortley 1990) by representing and estimating dyadic effects in a systemic context. Thus, our work moves toward the type of relational analyses Emirbayer (1997) discussed in his recent manifesto.

There are several methodological challenges to the estimation of the parameters in (1) that arise from the unusual structure of the data. First, the dependencies among actions between different pairs of actors can be accommodated by estimating parameters in the p^* framework, which employs maximization of the pseudolikelihood, conditioning the relation between each pair of actors (as indicated by the presence or absence of hostile action) on all other relations in the network.²⁰ Second, sets of dyads are nested within the blocks defined by the social structure as shown in table 1 (with additional blocks for relations involving at least one of the four singletons represented in fig. 2) and are therefore not independent.

The subgroup boundaries defining each block may become salient for actors' actions partly as actors generalize the transitivity of friendships. That is, actors tend to assume if i is tied to j and j is tied to k , then i is

²⁰ Our outcomes are atypical for the p^* or p^1 framework because our outcomes are not relations that are constant over time but indicators of specific action that occur at fixed points in time. Therefore, we might condition each outcome only on those indicators of action that were chronologically prior. But such conditioning would apply only to occurrences of hostile action, as there is no chronology associated with the absence of hostile action without full knowledge of each opportunity for hostile action. For example, we cannot know whether actor 3 abstained from hostile action against actor 23 prior to actor 11's abstention from hostile action against actor 7 unless we know the chronology of opportunities for hostile action, a level of detail not available in these data. Therefore, we specified our models in the standard p^* form. Through supplementary analyses, we established that our findings hold regardless of whether or not we conditioned on aspects of the pattern of hostile action included in the p^* portion of the model.

TABLE 2

TWO-LEVEL LOGIT MODEL OF HOSTILE ACTION BETWEEN A PAIR OF ACTORS (i, i')

FIXED EFFECT	COEFFICIENT		
	No Controls	Control for Subgroup Foci	Control for Subgroup Foci and Social Structure
Intercept	-3.712 (.222)	-3.835 (.193)	-3.826 (.201)
Members of the same subgroup	-1.287 (.553)	-1.198 (.564)	-1.229 (.688)
ρ^* control:			
Outstars090 (.184)	.110 (.188)	.101 (.187)
Transitive triads	1.000 (.225)	.998 (.254)	.985 (.312)
Same party318 (.313)	.308 (.276)
Both graduates of ENA		-.519 (.727)	-.491 (.798)
Friends018 (.677)
Two or more common friends			-.187 (.688)

NOTE.—The effect of each predictor has been centered around its grand mean so that β_0 roughly represents the overall tendency for actors to engage in hostile acts; SEs are in parentheses, $N = 756$ at level 1; 32 at level 2.

tied to k , even if they have not directly observed a tie between i and k (Freeman 1992b). This tendency to generalize is likely to be especially strong within cohesive subgroups that, by definition, contain an inordinate proportion of transitive triples (Davis 1967). If subgroup memberships are salient, then the pattern of hostilities in a given block may be cumulative or reflect an underlying characteristic of the block, such as when the members of different subgroups begin to perceive themselves as rivals (Cartwright and Zander 1968; Shaw 1981; Sherif and Sherif 1966). This pattern is consistent with the definition of the subgroups as moieties (Lévi-Strauss 1969).

Thus, we employ a two-level modeling framework that incorporates the random effects of blocks at the second level. Most important, the two-level framework allows us to define the effect of membership in the same subgroup at the level of the block and to assess the effect in terms of variance explained among the blocks. This represents an important response to concerns such as expressed by Davis (1994) regarding the limitation of logistic regression models in reporting percent variance explained. (The combination of ρ^* and multilevel models in estimating the regression coefficients and SEs is addressed more fully in the appendix.)

Estimates of parameters and approximate SEs for model (1) are given in the first column of table 2. The effect of subgroup membership is in the expected direction and is more than twice its approximate SE. The

coefficient of -1.287 indicates that the odds that an actor will engage in hostile action (vs. abstaining from hostile action) toward another are more than three and a half times greater if the two actors are members of the same subgroup than if they are not. This effect is observed controlling for the dependencies captured by the p^* model and quantifies the tendency observed in figure 2.

Of course, there are multiple systematic and idiosyncratic factors that might affect actors' actions. The capacity of these alternatives to account for the effect of subgroup membership can be explored by estimating the effect of subgroup membership while controlling for the alternatives. Because our focus is on the effect of subgroup membership on actors' actions, and because there are so few observed hostile actions from which to draw systematic inferences, we explore only those alternative explanations for hostile action that are likely to be confounded with subgroup membership. In particular, we first consider part of the bases for the foci of the subgroups—affiliation with the same party and graduation from ENA. These foci might affect action through bounded solidarity; members of a given party are commonly vulnerable when that party loses favor, and members of the intellectual elite are at times uniformly resented by French politicians (*Economist* 1995a).

We estimated and controlled for the effects of membership in the same party and attendance at ENA by entering dummy variables into our model (*same party* = 1 if actors i and i' were members of the same party, 0 otherwise; *both graduates of ENA* = 1 if actors i and i' were both graduates of ENA, 0 otherwise). In the second column of table 2, we note that the effects of common attendance at ENA and membership in the same party are small compared to their approximate SEs, but we retain them in the model because they are partly confounded with subgroup membership. Most important, the effect of subgroup membership is essentially sustained, indicating that the effect of subgroup membership cannot be attributed to commonalities associated with the measured subgroup foci.

The effects of subgroup membership also could be attributed to the structural conditions that form the basis of the subgroups. Members of a common cohesive subgroup are more likely to be friends and are more likely to have many mutual friends, other members of the subgroup, than other dyads in the system (Frank 1996; Mariolis 1982).²¹ It could be these

²¹ Mutual friendships also capture a form of structural equivalence among the actors (Burt 1991) and represent a characterization of triads of actors (Wasserman and Faust 1994). Supplemental analyses indicated that our assessment of our hypothesis was not affected when we controlled for other aspects of the friendship structure associated with the standard p_1 model (e.g., expansiveness, attractiveness) or Frank and Strauss's (1986) p^* triad model.

patterns of friendships defined at the dyadic and triadic level that account for tendencies in the patterns of hostile action currently attributed to subgroup membership. Therefore, we added a dummy variable indicating whether two actors were friends, and another variable indicating whether they shared two or more mutual friends, to our model.²² In theory, the effects of the structural conditions may be bidirectional. An actor may invest social capital in another because of the obligation of direct friendship or of mutual friends, or an actor may invest in another because he believes the other is likely to reciprocate, bound by a direct obligation of friendship or an indirect obligation through mutual friends.

Results for the model including control for the dyadic level social structure are given in the third column of table 2. As can be observed, the effects of friendships and common friends are minimal. Controlling for likely subgroup foci and the structural conditions that define the subgroup, the effect of membership in the same cohesive subgroup was sustained.²³ Thus, actors' actions align with subgroup boundaries more than with the specific social structure in which each dyad is embedded. Further, even if part of the explanatory power of subgroup membership could be attributed to covariates, in this case, a strong zero-order effect of subgroup membership (as is indicated in the first column of table 2) represents a parsimonious explanation for actors' actions. The subgroup effect encapsulates effects of both the subgroup foci and the social ties that form the bases of the subgroups.

The subgroup effect can be observed for members of subgroup *C* who compete for control of Grande Banque in a highly competitive international marketplace but who, as far as we know, have refrained from hostile actions against each other. It also applies to the actors in subgroup *A* who tend to refrain from hostile action against each other while they compete for government posts and resources in a turbulent political climate. In fact, even when the friendship between subgroup *A* members 12 and 24 was dissolved during actor 24's service in the cabinet of a conservative prime minister, there were no reports of hostile actions between the two.

²² We use two or more friends because preliminary analyses revealed that this effect was stronger and more interpretable than effects associated with one or more mutual friends or three or more mutual friends.

²³ We note that the model does not indicate the unadjusted effects of the covariates on hostile action, many of which may be greater to the extent that the reported effect is reduced due to correlation with membership in the same subgroup (e.g., direct friendship and membership in the same subgroup are, of course, highly correlated). But our focus is on the effect of membership in the same subgroup, controlling for the covariates, and these effects are adequately represented in our models. Indeed, the power of membership in the same subgroup is underscored by the fact that it is still an important predictor even after controlling for the covariates.

Though not definitive, we take the evidence presented here as quite consistent with our first hypothesis, that actors are more likely to enforce each other's trust if they are members of the same subgroup than if they are not members of the same subgroup. In the next section, we establish the negative effect of subgroup membership on supportive action (indicative of reciprocity transactions) and then in the following section we contrast the two subgroup effects.

Supportive Action and Reciprocity Transactions

We begin by establishing that there is a norm of reciprocity among the French financial elite, as there is in most systems in which transactions are embedded in even a moderately dense social structure (Coleman 1988; Dawes 1988; Portes and Sensenbrenner 1993). A norm of reciprocity should allow actors to invest in one another with the belief that the other will feel obligated to reciprocate, thus facilitating a fluid movement of resources throughout the system. First, actors have indicated that they sometimes support one another to return a favor. For example, actor 5, explaining that his institution did not take a profitable action that would have resulted in an enormous loss for an actor on the margin of the French financial elite, said, "If we didn't have a debt of gratitude to [the marginal actor] we would have sold" (De Quillacq 1992, p. 124). Further, negative sanctions have been imposed on actors who failed to reciprocate, which would explain why no one has supported actor 3 after he failed to reciprocate for actor 20's support in a raid on actor 5's institution.

In order to operationalize the pursuit of social capital through reciprocity transactions, we identified instances of supportive action, defined as those actions that helped another maintain or acquire further resources or that helped another to retain or obtain a high-level position (which likely affects the supported actor's access to resources).²⁴ As examples, actor 27 purchased extra shares in actor 5's institution to support the stock price and rebuff a raid initiated by actor 3, and actor 1 recommended his friend, actor 28, for a high-level position in a large bank.²⁵

As was the case regarding enforceable trust, our ability to directly measure the accumulation of social capital through reciprocity transactions is limited in our data because we do not have adequate longitudinal information and indicators of actors' states of mind. But given the established theory and evidence regarding tit-for-tat exchanges (e.g., Coleman 1988;

²⁴ Many of these actors can exchange positions because they have comparable cultural capital, such as obtained through graduation from ENA (Bourdieu 1986).

²⁵ As was the case for hostile action, we note that these actions do not necessarily represent enduring relations but constitute large and public one-time actions

Dawes 1988; Portes and Sensenbrenner 1993) and examples of a norm of reciprocity among the French financial elite, we take supportive actions as a rough indicator of social capital pursued through reciprocity transactions. Further, we focus on acts of support and not on reciprocation because it is the act of support that generates social capital as it creates an obligation in the other. In fact, when an act of support is adequately reciprocated, the actor who offered the initial support no longer has social capital by our definition.

In figure 3, we represent supportive actions within the geometric space of figure 1. There are two critical features to note in juxtaposing figure 3 with figures 1 and 2. First, no actor has acted hostilely against another who had previously supported him. This is consistent with our claim that a large investment of social capital through a supportive act not only is likely to elicit reciprocity but also reduces the likelihood of hostile action. Second, there are considerable supportive acts spanning between subgroups. The trend observed in the graphic is consistent with a rephrasing of hypothesis 2 in terms of the relationship between supportive action and subgroup membership among the French financial elite:

RESTATEMENT OF HYPOTHESIS 2.—*Ceteris paribus, among the French financial elite, actors are more likely to act supportively (support being consistent with reciprocity transactions) if they are not members of the same subgroup than if they are members of the same subgroup.*

To ascertain the effect of membership in the same cohesive subgroup on the occurrence of a supportive act, we estimate parameters in a model similar to (1), except with the dependent variable being the occurrence or lack of occurrence of an act of support instead of an act of hostility.²⁶ As was the case for our model of hostile action, we control for effects associated with p^* models and consider alternative explanations associated with subgroup foci and aspects of the social structure.²⁷ Again, we argue that

²⁶ We removed the mutually supportive actions between 1 and 8 because they did not represent evidence of social capital, as they were implemented simultaneously, and thus neither actor was reliant upon a norm of reciprocity to insure reciprocity.

²⁷ Optimally, we would control for the expected economic gain associated with a given act of support, but measuring such a covariate is problematic. First, it is unclear how accurately the actors themselves are able to assess the potential economic gain of a supportive act. Second, it would be difficult to identify opportunities for supportive action *not* undertaken, let alone to assess the potential gain associated with those opportunities. Moreover, we argue that the potential economic gain is determined largely by factors associated with the particular actors and their circumstances and is not likely to be confounded with subgroup membership, other than through effects associated with the subgroups. That is, the extent to which the return on a supportive act varies according to the *social structure* is most accurately attributed to dyadic and subgroup effects that we include in our models and not to unmeasured attributes of *individuals*.

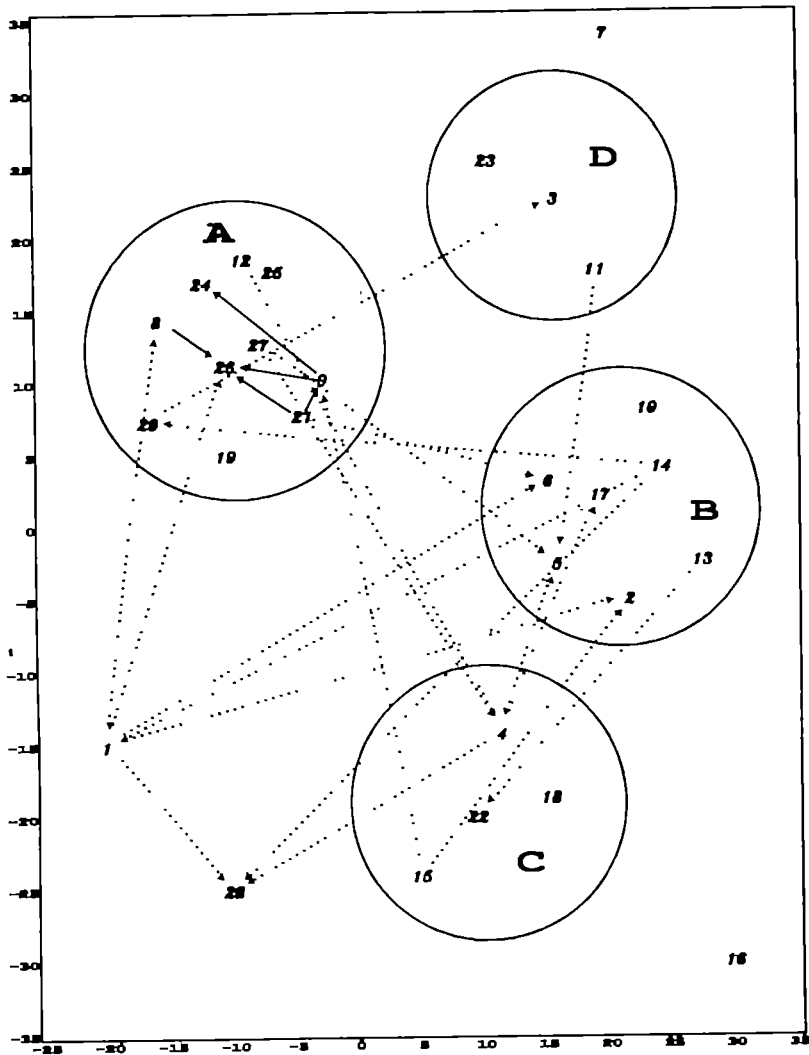


FIG. 3.—Supportive acts embedded in the friendship structure (solid lines indicate acts within subgroups, and dotted lines, acts between subgroups).

the effects of the covariates may be bidirectional. An actor may support a member of his political party, fellow alumnus, friend, or friend of a friend because of a socially based bond (such as through bounded solidarity or "value introjection") or because he believes that such others are more likely to reciprocate.

The estimates for the model of supportive action are presented in table

TABLE 3
TWO-LEVEL LOGIT MODEL OF SUPPORTIVE ACTION

FIXED EFFECT	COEFFICIENT		
	No Controls	Control for Subgroup Foci	Control for Subgroup Foci and Social Structure
Intercept	-3.576 (.211)	-3.631 (.199)	-3.781 (.231)
Members in the same subgroup	-.365 (.593)	-.265 (.598)	-1.104 (.727)
β^* control			
Outstars263 (.120)	.288 (.116)	.272 (.114)
Transitive triads	1.275 (.250)	1.305 (.234)	1.313 (.219)
Same party		-.939 (.530)	-1.129 (.501)
Both graduates of ENA		-.014 (.264)	-.191 (.285)
Friends			1.665 (.346)
Two or more common friends			-.151 (.659)

NOTE.—The effect of each predictor has been centered around its grand mean so that β_0 roughly represents the overall tendency for actors to engage in hostile acts; SEs are in parentheses, $N = 756$ at level 1, 32 at level 2.

3. The effects of membership in the same party and common graduation from ENA are negative (although the latter is small relative to its approximate SE). As indicated in the third column of table 3, the effect of friendship is strong and positive, perhaps because actors believe that a friend is likely to fully reciprocate an act of support or because of value introjection—actors simply value the social tie more than the economic gain or loss associated with the act of support. Controlling for the effects of direct friendship and membership in a common subgroup, the effect of mutual friends is minimal.

Most important, the coefficient of -1.104 in the third column of table 3 indicates that the odds that an actor will at one time support (vs. never supporting) a nonsubgroup member are roughly three times greater than the odds that he will support a subgroup member. This effect is observed controlling for dyadic effects of social structure and institutional commonalities. In comparing the second and third columns, we note that the subgroup effect increased in magnitude as a result of controlling for components of the social structure. (The pattern of effects across the columns is similar for common affiliation with a political or educational institution, but the changes are less dramatic relative to their SEs.) Some of the negative effect of affiliation with a common subgroup on supportive action is mitigated by the friendships that one is likely to form and share with others affiliated with the institution or subgroup. For example, perhaps

the three actors within subgroup *A* (8, 9, and 21) who supported actor 26 would have been less likely to do so had they not been direct friends of actor 26. But in general, our findings run counter to expectations based on group solidarity. Once controlling for aspects of social structure defined at the level of the dyad and triad, actors are *less* likely to support a member of their specific old boy network than others in the system at large.²⁴

We observe the effect of subgroup membership in the pattern of absence of support among members of subgroup *B*. Although many of these men hold positions in the banking industry and are aware of each other's plights and opportunities, there is no example of a member of subgroup *B* definitively supporting another member of subgroup *B*. Instead, for example, actor 2 relied on the support of actors 1 and 15 to ascend to the head of Grand Banque. Similarly, actor 14 recommended actors 1, 20, and 28 for positions or partnership in a major investment bank but has not done the same for any member of his subgroup.

Abstinence from Hostile Action within Subgroups versus Supportive Action outside Subgroups

As we have shown among the French financial elite, action is linked to the social structure delineated by subgroup boundaries as actors are less likely to engage in hostile action within their subgroups and more likely to support others not in their subgroups. Now we estimate a single model of "stacked" data (Fienberg, Meyer, and Wasserman 1985) so that we may contrast the positive effect (1.229) of subgroup membership on abstinence from hostile action reported in the third column of table 2 with the negative effect (-1.104) of subgroup membership on supportive action reported in the third column of table 3. The contrast is defined by the interaction of subgroup membership and type of action in predicting whether or not we observe evidence of the action. Formally, the contrast is a test of the following hypothesis:

RESTATEMENT OF HYPOTHESIS 3.—*Among the French financial elite, actors are more likely to abstain from hostile action (consistent with enforced trust) versus act supportively (consistent with reciprocity transactions) if they are members of the same subgroup than if they are not members of the same subgroup.*

²⁴ The effect of subgroup membership on supportive action could reflect our nonrandom sample of data, through which we have become more aware of supportive action between subgroups than within them. We respond by turning our attention to those acts of support that were most likely to be publicly known (because of their large commitment of resources) and are therefore less likely to be biased by any filter. Such acts occurred exclusively between members of different subgroups or involved unaffiliated actors, thus supporting the findings we report.

The linkage of the two types of action reflects the fact that there is a single process, the pursuit of social capital, underlying both types of action. Thus, while not mutually exclusive, the two types of action tend to occur in different regions of the social structure.

We do not report all of the estimates for this combined model, which essentially reproduce the findings in tables 2 and 3. The estimate for the contrast of subgroup effects was 2.100 and was more than three times its approximate SE of .619.²⁹ A regression using the quadratic assignment procedure (QAP) based on the level-2 model indicated $P \leq .015$ from 1,000 nonparametric monte carlo simulations.³⁰ The estimate of the contrast of subgroup effects was 1.35 with approximate SE .43 before controlling for covariates.

Controlling for covariates, the effect of type of action (abstaining from hostile action over active support) on the odds that an actor engaged in the action were eight times greater if the other was a subgroup member than if he was not. That is, relative to other pairs of actors in the system, there is a strong tendency for members of subgroups to abstain from hostile action against each other rather than to actively support one another. Further, membership in a common subgroup explained 38% of the variation across blocks defining the second level of our models.

The effect of the subgroup boundaries is typical in subgroup *A* in which members never engaged in hostile action against one another but also rarely supported subgroup members who were not friends. By abstaining from hostile actions within the subgroup, they sustain the solidarity of the friendship structure (many of the friendships likely were formed or strengthened through common association with ENA and the treasury). Able to enforce the trust of members of their subgroup, they turn to others outside the subgroup, such as the unaffiliated actor 1 and members of subgroups *B* and *C*, for access to resources that members of their own subgroup are unlikely to provide.

²⁹ The estimate for the contrast of subgroup effects after controlling for covariates was 2.06 using the single-level p^* approach described by Wasserman and Pattison (1996) and was associated with a change in deviance based on the pseudolikelihood of 4.0 (which can roughly be compared to a χ^2 distribution on 1 df —in this case suggesting $P \leq .05$).

³⁰ The regression was calculated by first estimating a multilevel model with just the covariates and a dummy variable for type of action as independent variables at level 1. The empirical Bayes residuals of the effect of type of action were then saved (the empirical Bayes residuals adjust for the different numbers of pairs within blocks—see Bryk and Raudenbush 1992). QAP (Hubert 1987) in UCINET IV (Borgatti et al. 1992) was then used to estimate the level-2 regression of these residuals on a dummy variable indicating whether the block represented relations within a single subgroup or not.

Conversely, the effect of lack of subgroup membership is most pronounced for actor 1. Actor 1 did not have the opportunity to establish a host of friendships through attendance at ENA nor through service in a formal government ministry and is not affiliated with a particular subgroup. Unconstrained by a dense set of friendships through which his trust might be enforced, actor 1 has engaged in hostile action toward members of each of the three most active subgroups (*A*, *B*, and *C*). Unable to enforce the trust of others, he has actively supported, and relied on the support of, a host of others in subgroups *A* and *B*. Further, at the level of the individual actor, the separate types of acts are interdependent, because acts of support toward one actor often imply an act of hostility against another. For example, when actor 1 supported actor 2 for the head of Grande Banque, he effectively undermined the ambitions of actor 9 for the same position. Thus, actor 1 has achieved his status through actions that are suited to his low level of social affiliation. This has allowed him to frequently act on the advantage he gains through his close ties with recent prime ministers.

SUMMARY AND DISCUSSION

A major theme of the resurgent area of economic sociology is that the nature of a transaction, even between two actors who compete for economic gain, is affected by the general social, cultural, and institutional contexts in which it is embedded (Smelser and Swedberg 1994; Zafirovski and Levine 1997). In this article, we extended Granovetter's (1985) emphasis on economic action as embedded in a *social* context (see also Baker 1989; White 1981) by hypothesizing that social contexts, and therefore types of action, vary *within* systems. Specifically, we drew on developing theory (in particular that of Portes and Sensenbrenner 1993) to argue that actors are more likely to pursue social capital through enforceable trust when their action is embedded in the dense social ties found within cohesive subgroups than when it is not.

Where ties are more sparse, actors are more likely to pursue social capital through reciprocity transactions, supporting each other and then relying on systemic norms to help insure reciprocation. This norm of delayed reciprocation amidst sparse ties is not easily explained by theories of group solidarity. Instead, we argue that actors generally pursue social capital but that the mechanism, enforceable trust or reciprocity transactions, depends on the social structure in which action is embedded. Outside of subgroup boundaries, where it is difficult to enforce trust, actors rely on reciprocity.

In order to explore our hypotheses, we examined the system of

the French financial elite. We began by representing variation in the social structure as a dual organization in terms of friendships within and between cohesive subgroups, confirming the empirical tendency for friendships to be concentrated within the subgroup boundaries. We then presented anecdotal evidence of sanctions for violating the general norm of solidarity and the specific norm of subgroup solidarity. We also described instances in which actors perceived the necessity to reciprocate an earlier act of support as well as sanctions imposed for actors who did not reciprocate the support of others. These instances are indicative of the general norms that must be present for actors to accumulate social capital.

To describe the relationship between social structure and action, we linked the tendencies for hostile action (the lack of which is consistent with enforceable trust) and supportive action (the presence of which is consistent with reciprocity transactions) with the subgroup structure through graphical displays. Indeed, it is quite striking how much the seemingly unpredictable pattern of directed actions among the French financial elite aligned with subgroup membership. Thus, we observe the paradoxical tendency for action between members of different subgroups (and involving unaffiliated actors) to be hostile *and* supportive, knitting the moieties into a system with a dual organization.

Through our multilevel models, we separated subgroup effects that differentiate among blocks (e.g., membership in the same subgroup or not) from effects that differentiate pairs of actors (e.g., friendship patterns and affiliation with political parties and educational institutions). Estimates of subgroup effects on hostile and supportive action quantified and confirmed the tendencies we observed in the juxtaposition of our figures. Ultimately, by combining results across hostile and supportive action, we demonstrated that action between subgroups was different from that within subgroups. Though actors in general abstain from hostile action more frequently than they engage in directly supportive action, this tendency is more pronounced between members of the same subgroup than between other pairs of actors. Thus, even for these actors where identities are closely tied to the resources they manage, the social structure matters.

In any system, even if the variation explained by subgroup membership could be attributed to multiple factors at the level of the dyad and triad, the subgroup consolidates these factors into a single supradyadic entity that facilitates characterization of the contexts for action and systemic structure. But among the French financial elite, subgroups do more than simply consolidate other effects. We found that subgroup membership predicted nearly 40% of the variation between blocks even after controlling for the constitutive characteristics of subgroups. That is, our evidence is consistent with the characterization of the subgroup as an entity (Abbott

1996), drawing coherence from the pattern of friendship and having a direct effect on actors' actions.¹¹

By studying a small system of the French financial elite, we observe how the extensive consolidation of educational experiences, social structure, political affiliation and government participation generates a microcosm, reflecting the most basic French institutions. In particular, we characterize the system of French financial elite in terms of a dual organization, featuring cohesive subgroups in which friendships are concentrated but between which important transactions occur.

Even if the specific membership of the French financial elite changes, we would anticipate that the dual organization would persist as long as two critical conditions hold: First, access to the elite must be established through filtering institutions that cultivate social ties. Second, access to resources is partially dependent on a fluctuating political system, thereby giving members of the French financial elite opportunity for alliances and quick action that often involves others not in their immediate social circle. The social ties help to establish the structure through which norms of solidarity and reciprocity are generated and enforced. The fluctuating political system establishes the basis and benefit of crosscutting actions that draw on, and perpetuate, the norms. Though these conditions are considerable, recall that Lévi-Strauss (1969) argued that the dual organizations, such as found in the French financial elite, occur in many systems.

While our characterization might be valuable to those who seek to engage in profitable transactions with members of the French financial elite (consistent with one of the goals of De Quillacq [1992, p. ix]), our methods could more generally be used to help those who intend to understand and cultivate social capital within a system. If, as Woolcock (1998) argues, attempts to build social capital must access existing social structures, our methods can help reveal those structures and how they are linked to action by characterizing the existing distribution of social capital.

Our orientation may also be applied to the study of other sets of elites where we may look for a dual organization. More generally, we may look for an alignment between exogenous cultural institutions and internal social structure and dynamics. Perhaps this will help us to understand under what circumstances elites reproduce versus challenge existing institutions.

In terms of theory, we characterize the subgroups as mesolevel entities.

¹¹ The subgroup effect may be due to actors' tendencies to perceive a generalization of the transitivity of friendships among members of a subgroup. That is, actors may construct subgroup meaning through their sensemaking (Bourdieu 1984; Weick 1993), beyond the instrumental effects of the observed ties that define the subgroup structure.

As such, they structure the transition from microlevel actions to the macrolevel phenomenon of systemic coherence. Subgroups are sustained through frequent, small, and passive investments of enforceable trust, while a system is strengthened by a dual organization as members of different subgroups and unaffiliated actors occasionally engage in the large, discrete investments of reciprocity transactions (in the terms of Woolcock [1998] systems need both "linkage" and "integration").

Of course, there will be some instances when our theory and hypotheses will not apply. In particular, actors may not be able to enforce the trust of subgroup members when there is little effect of sanctions imposed for the violation of subgroup solidarity. This may occur when the underlying social ties are unstable or sparse, such as in systems comprised of the leaders of nations or corporations or in a fully competitive marketplace of strangers. In these cases, the internal social structure of the system is relatively undifferentiated. The actors may still accumulate social capital independently through reciprocity transactions but not through allegiance to a web of ties such as those within a cohesive subgroup.

In other cases, reciprocity transactions may not be distributed throughout the system. When a system becomes factionalized (as opposed to the partially integrated subgroups of a dual organization), support for a member of another subgroup may be taken, by definition, as hostile to one's own subgroup (e.g., Sherif and Sherif 1966). This may occur when divisions between subgroups become institutionalized or formally designated, such as through gang membership or team affiliation. In these cases, the system degenerates into a set of nonintegrated subsystems, each defining a faction.

Even when our theory regarding the link between social capital and subgroup membership does not apply, our methods can be generalized to inform the study of other systems of actors. The classical literature describing many systems as comprised of nonoverlapping cohesive subgroups (e.g., Blau 1977; Festinger et al. 1950; Homans 1950; Lévi-Strauss 1969; Nadel 1957; Simon 1965) motivates us to seek to identify and determine the salience of cohesive subgroups in the pattern of ties among other sets of actors. This pattern of ties may establish an underlying structural integrity of many systems and is thus consistent with our representation in a "crystallized" sociogram. Our graphical techniques can also be used to overlay patterns of action and social structure to reveal the specific contexts of actors and to integrate multiple forms of data. Ultimately, as we have done here, p^* models can be extended to test propositions regarding the relationship between action and social structure by specifying effects at the level of the pair and block.

Finally, perhaps in other instances it will be possible to focus more on

the *processes* through which action and social structure become interrelated. In our example, though the presence of social capital provides a plausible explanation for our findings, we acknowledge that we have not directly observed the processes through which actors accumulate social capital. To our advantage, our theoretical link draws on the conditions of enduring ties more than on the specific conditions of the social structure at the time of a given action. Thus, our analyses and explanations are conducted almost independent of a systemic chronology. But we echo Friedkin's (1997) call for longitudinal data and corresponding methods that would allow us to more directly study the processes through which action and social structure become interrelated.

APPENDIX

There are two recent advancements in the estimation of models such as (1) that have important implications for specification and substantive interpretation. First, in obtaining estimates of the parameters one might be inclined to use maximum likelihood techniques such as are available to estimate the parameters in logit models (e.g., Agresti 1984). But it is difficult to define the likelihood of the parameters in models such as (1) as a function of the data because the observations are not independent. To begin, an act of hostility from i to i' is not independent of an act of hostility from i' to i . Such dependencies are accounted for in the standard p_1 model (Fienberg and Wasserman 1981; Fienberg et al. 1985; Holland and Leinhardt 1981) by specifying the set of relations among the dyad as the unit of analysis (including the relation between i and i' as well as the relation between i' and i). While the p_1 approach represents an important advancement in the estimation of parameters such as in model (1), the respecification of the models still does not account for dependencies among pairs outside the dyadic unit.

But a new estimation approach, developed by Frank and Strauss and their colleagues (Frank 1991; Frank and Nowicki 1993; Frank and Strauss 1986; Strauss and Ikeda 1990) and described by Wasserman and Pattison (1996) shows that the estimation procedure used for standard logit models can be used to obtain estimates based on maximization of the "*pseudolikelihood*." The pseudolikelihood conditions the relation between each pair of actors on the relation between *every other pair of actors in the network*. In order to condition on every other relation, one must respecify the model to control for aspects of relations between dyads and triads of actors. For example, one might condition on whether or not a relation is reciprocated, or whether actors i and i' have a common relation with another (one only need condition on n characteristics of the network structure assuming one

does not wish to estimate unique characteristics of each actor—see Wasserman and Pattison 1996, p. 410). The p^* model for the hostile relation of model (1) can be written as

$$\log \left(\frac{P[HA_{ii'} = 1 | HA_{ii'}^c]}{1 - P[HA_{ii'} = 1 | HA_{ii'}^c]} \right) = \beta_0 + \beta_1(MMS_{ii'}) + \theta'V_{ii'}, \quad (A1)$$

where $HA_{ii'}^c$ is the complement of $HA_{ii'}$, representing the presence or absence of hostile action between all other pairs of actors. MMS denotes members of the same subgroup, and V denotes vector of characteristics capturing hostile (or absence of hostile) acts in the network.

In order to determine the vector of network statistics on which to condition the outcome in (A1), we began by estimating terms in the equivalent of the standard p_1 model, that is, controlling for individual expansiveness and attractiveness and reciprocity (because reciprocity represents the effect of the action of i' toward i on the action of i toward i' , reciprocity was coded according to chronological sequence). Using differences in pseudolikelihood ratio statistics, we determined that the sets of expansiveness and attractiveness parameters did not improve the fit of the model (using a liberal χ^2 based cutoff of $P \leq .10$). While the reciprocity parameter did improve the fit, we opted instead to explore a set of homogeneous effects (those that are not specific to individual actors) associated with the *triad model* as described by Frank and Strauss (1986) and as calculated using the p^* software (Crouch and Wasserman 1998).¹² Again, using a liberal χ^2 cutoff associated with $P \leq .10$, we retained the parameter for two outstars (essentially the number of hostile acts that actor i commits). We also chose to retain the parameter for transitive triads (the number of common others against whom both i and i' acted hostilely—the only positive occurrences of this involved the triad of actors 1, 9, and 18) in order to maintain compatibility with our model for support even though they did not greatly reduce the pseudolikelihood ratio statistics. These characteristics then represent the aspects of the set of hostile actions upon which we conditioned when we estimated the p^* models reported in our text. In the model of supportive acts, we also conditioned on the outstars and transitive triads. Thus, the vector of network characteristics in (A1) included outstars and transitive triads, with corresponding estimates for

¹² The reciprocity predictor took a positive value in only one case for hostile action and thus essentially defined an indicator for the one case of reciprocity. Therefore, we did not consider it worthwhile to estimate the parameter. We verified that our choice of terms on which to condition in the p^* model did not affect our conclusions regarding our hypotheses.

θ_{outstars} and $\theta_{\text{transitive triads}}$ reported in tables 2 and 3. In the combined model, we conditioned on outstars and transitive triads, as well as the interaction of outstars and transitive triads with the mechanism for pursuing social capital in order to preserve the conditioning of the p^* model.

Before estimating the parameters in model (A1), we respecified the model based on the structure of table 2 and figure 1. In particular, the pairs of actors are nested within blocks characterized by the respective subgroup components. For example, the pair involving actors 4 and 26 is nested within the block defined by subgroups *C* and *A*. The nature of the dependencies among pairs of actors within blocks suggests a two-level model should be specified and estimated (Bryk and Raudenbush 1992). The two-level model allowed us to assess the effect of our most critical independent variable, membership in the same subgroup, in terms of variation explained at the block level (models such as developed by Fienberg et al. [1985] incorporate block-level effects as fixed effects, requiring estimation of the unique effect of each block, and do not provide an estimate of the variance among blocks against which the effect of membership in the same subgroup can be compared). Therefore, we defined n pairs of actors (there are a total of $28 \times 27 = 756$ pairs) nested within J blocks (*AA*, *AB*, . . . , *IH*—there were 32 blocks, based on the four sets of relations within subgroups and the 28 sets of relations between the four subgroups and the four unaffiliated actors). Using ii' to index the pairs in block i , we estimated the parameters in a two-level model.

At level 1:

$$\log \left(\frac{P[\text{HA}_{ii'} = 1 | \text{HA}_{ii'}^c]}{1 - P[\text{HA}_{ii'} = 1 | \text{HA}_{ii'}^c]} \right) = \beta_{ij} + \theta_{\text{outstars}, \text{outstars}_{ii'}} + \theta_{\text{transitive triads}, \text{transitive triads}_{ii'}} \quad (\text{A2})$$

at level 2:

$$\beta_{ij} = \gamma_{00} + \gamma_{01} \text{MSS}_j + u_{0j},$$

$$\theta_{\text{outstars}, j} = \gamma_{40},$$

$$\theta_{\text{transitive triads}, j} = \gamma_{50}.$$

Conceptually, within each block j , a unique estimate of each level-1 coefficient (β_{ij} , $\theta_{\text{outstars}, j}$, $\theta_{\text{transitive triads}, j}$ —we omit the level-1 coefficient β_{ij} , which was originally associated with the effect of membership in the same subgroup) is obtained. These coefficients then become outcomes in the level-2 model. In this case, the regression coefficients for the covariates (including the set of p^* characteristics) were defined to be constant across the

blocks, thus forming an equivalence in terms of the covariates between models (A1) and (A2). We then defined and estimated the effect of membership in the same subgroup at level 2, depending on whether the actors were members of the same cohesive subgroup (blocks *AA*, *BB*, *CC*, or *DD*) or not. The random effects at level 2 capture the fact that patterns of directed actions vary within each subgroup (Hechter 1983) and between subgroups. Our focus is then on the ability of membership in the same subgroup to explain variation across the level-2 blocks.

Estimates of hierarchical generalized linear models such as (A2) can be obtained via maximization of the penalized quasilielihood (based on a second order Taylor approximation of the likelihood—see Breslow and Clayton [1993]; Goldstein [1991]; or Raudenbush [1995] for a discussion of multilevel generalized linear models). We used the software developed by Bryk, Raudenbush, and Congdon (1997) to obtain estimates of the variance components and empirical Bayes estimates of fixed effects (although we have confirmed that each estimate is within .15 of those obtained using software developed by Hedeker and Gibbons 1996). To apply this approach to models for social network data such as the p^* models, we begin by writing the log likelihood of the parameters in a two-level model as the joint density function of Y (the vector of dependent variables) given a vector of fixed effects, Θ_1 , and a vector of random effects, Θ_2 (see Bryk and Raudenbush 1992 or Raudenbush 1995):

$$\log \text{likelihood}(\Theta_1, \Theta_2) = \log f(Y|\Theta_1, \Theta_2). \quad (\text{A3})$$

For logit models at level 1, Raudenbush (1995) writes the likelihood as:

$$\log f(Y|\Theta_1, \Theta_2) = Y'(A_1\Theta_1 + A_2\Theta_2) - \sum_{j=1}^J \sum_{i=1}^{n_j} \log(1 - p[y_{ij} = 1]), \quad (\text{A4})$$

where A_1 is a vector of predictors associated with the fixed effects, A_2 is a vector of predictors associated with the random effects, and $p[y_{ij} = 1]$ represents the probability that the outcome takes a value of one for actor i in subgroup j . For model (A2) the outcome represents hostile action directed from i to i' within block j : $HA_{ii'j}$. In order to write the pseudolikelihood, we condition the outcome not only on the parameters contained in (A4) but also on values of the matrix of relations other than the specific relation between i and i' in block j ($HA_{ii'j}^*$). Therefore, we add to the model Z , the vector of network statistics on which the p^* model is conditioned (in our example, Z consists of outstars and transitive triads), as well as the corresponding vector Θ_3 (in our example Θ_3 consists of θ_{outstars} and $\theta_{\text{transitive triads}}$):

$$\sum_{i=1}^{n_j} \sum_{i'=1}^{n_j-1} \log f(\text{HA}_{ij} | \Theta_1, \Theta_2, \Theta_3, \text{HA}_{ij}^c) = \quad (\text{A5})$$

$$H'(A_1\Theta_1 + A_2\Theta_2 + Z\Theta_3) - \sum_{j=1}^J \sum_{i=1}^{n_j} \log(1 - p[\text{HA}_{ij} = 1]),$$

where H' is the matrix of the dependent variables, HA_{ij} , stacked into a single vector.

Because maximum likelihood estimation of logit models can be extended to maximization of the pseudolikelihood for social network models (Frank and Strauss 1986; Strauss and Ikeda 1990; Wasserman and Pattison 1996), we reason that an algorithm that maximizes the penalized quasilielihood for logit models with random effects as in (A4) then maximizes the penalized quasi-pseudolikelihood for social network logit models with random effects as in (A5). Of course, one could directly obtain pseudolikelihood estimates by ignoring the hierarchical nature of the data and not including the random effects in level 2, and we report the results of such estimates for our final model.

The specification of our models in the multilevel framework also suggests a potential approach for estimating SEs of effects defined at level 2,¹³ such as the effect of membership in a common subgroup, which can be assessed essentially against the variation among the level-2 blocks, and is therefore not strongly affected by dependencies among observations *within* blocks. This respecification also reveals an important limit in our power to test for the effect of membership in a cohesive subgroup. There are at most only 31 degrees of freedom (based on the 32 level-2 units) for assessing the effect of any predictor defined at level 2 of our models.

Even if we had more units at level 2, the calculation of SEs assumes that the errors at level 2 are normally distributed. Analyses of residuals from each of our models (see Bryk and Raudenbush 1992, chap. 9) indicated that the distribution of residuals deviated from their comparison distribution (e.g., for the models in the third column of tables 2 and 3, the residuals were positively skewed because hostile and supportive acts were relatively uncommon in most of the units, and in the final model the residuals associated with the contrast of the different mechanisms for pursuing

¹³ As Burt (1991) and Wasserman and Faust (1994) have noted, the issue of significance testing in social network analysis requires serious consideration as one typically has data from the whole population rather than from a sample. We interpret our significance tests as an evaluation of phenomena sampled from a fixed time interval relative to a population of all phenomena over the duration of the system.

social capital also were skewed). Therefore, we report the "robust" SEs as described by Raudenbush (1995) and Zeger, Liang, and Albert (1988).²⁴ Further, the reported SEs are not robust with respect to the assumption that the level-2 units are independent, that is, the extent of hostile actions between members of subgroups *A* and *B* is independent of the extent of hostile actions between members of subgroups *A* and *C*. This assumption may not be tenable to the extent that actions in the entire system are dependent, although it may be more likely to be satisfied than the assumption in the single-level framework in which even actions between members of the same subgroup are assumed to be independent. Finally, the SEs are consistent only as the number of level-2 units grows large. Because of the concern regarding the dependencies among the level-2 units (blocks) and consistency of the robust SEs in only 32 blocks, we use the SEs only as a guideline for interpreting our effects, and we emphasize the size of the effects and variance explained over interpretation of the *P* values.

While our use of the approximate SEs is tentative, we know of no clearly superior way of assessing the significance of our estimated coefficients. Wasserman and Pattison (1996) compare "likelihood ratio statistics" (although the likelihood is not well understood) from nested models to assess the value of including given terms in a model. Although we follow their practice in note 29 above, little is known about the properties of these deviance statistics for *p** models. Therefore, in this instance, we believe that it is more valuable to report an estimate defined at level 2 of our models and assessed relative to variation explained at level 2 than to assess the effect primarily in terms of a comparison of likelihood ratio statistics that do not directly incorporate the multilevel nature of the data. Further, we confirm our findings by reporting a *P* value based on a QAP regression (Hubert [1987], as implemented in Borgatti, Everett, and Freeman [1992]) at level 2 of our last model.

REFERENCES

- Abbott, Andrew. 1996. "Things of Boundaries." *Social Research* 62:857-82.
 Agresti, A. 1984. *Analysis of Categorical Data*. New York: Wiley & Sons.
 Baker, Wayne E. 1989. "Social Structure of a National Securities Market." *American Journal of Sociology* 89:775-811.
 Barnes, J. 1972. "Social Networks." Pp. 1-29 in *Anthropology Module 26*. Reading, Mass.: Addison Wesley.

²⁴ In spite of their name, the "robust" SEs are not necessarily smaller than the SEs estimated based on distributional assumptions. For the contrast of subgroup effects on hostile and supportive acts the nonrobust SE was approximately 1.19, and the *t*-ratio based on the comparison of the parameter estimate to this SE was 1.76, with *P* approximately $\leq .09$.

- Bian, Yanjie. 1997. "Bringing Strong Ties Back In: Indirect Ties, Network Bridges, and Job Searches in China." *American Sociological Review* 62:366–85.
- Blau, P. M. 1977. *Inequality and Heterogeneity*. New York: Macmillan.
- Borgatti, P., M. Everett, and L. Freeman. 1992. *UCINET IV, Version 1.0 Reference Manual*. Natick, Mass.: Analytic Technologies.
- Bott, E. 1971. *Family and Social Network: Roles, Norms and External Relations*. London: Tavistock.
- Bourdieu, P. 1984. *Distinction: A Social Critique of the Judgement of Taste*. Cambridge, Mass.: Harvard University Press.
- . 1986. "The Forms of Capital." Pp. 241–58 in *Handbook of Theory and Research for the Sociology of Education*, edited by J. G. Richardson. New York: Greenwood Press.
- Breslow, N., and D. Clayton. 1993. "Approximate Inference in Generalized Linear Mixed Models." *Journal of American Statistical Association* 88:9–25.
- Bryk, A., and S. Raudenbush. 1992. *Hierarchical Linear Models*. Newbury Park, Calif.: Sage Publications.
- Bryk, A., S. Raudenbush, and R. Congdon. 1997. *Hierarchical Linear Models*. Chicago: Scientific Software.
- Burt, R. S. 1982. *Toward a Structural Theory of Action*. New York: Academic Press.
- . 1991. *Structure*. New York: Columbia University.
- . 1992. *Structural Holes*. Cambridge, Mass.: Harvard University Press.
- Cartwright, D., and F. Harary. 1956. "Structural Balance: A Generalization of Heider's Theory." *Psychological Review* 63:277–93.
- Cartwright, D., and A. Zander. 1968. *Group Dynamics: Research and Theory*. New York: Harper & Row.
- Coleman, James S. 1988. "Social Capital in the Creation of Human Capital." *American Journal of Sociology* 94 (suppl.): S95–S120.
- . 1990. *Foundations of Social Theory*. Cambridge, Mass.: Harvard University Press.
- Cooley, C. H. 1909. *Social Organization*. New York: Schocken.
- Crouch, Brad, and Stanley Wasserman. 1998. "A Practical Guide to Fitting p* Social Network Models via Logistic Regression." *Connections* 21 (1): 87–101.
- Davis, J. A. 1967. "Clustering and Structural Balance in Graphs." *Human Relations* 20:181–87.
- . 1994. "What's Wrong with Sociology?" *Sociological Forum* 9 (2): 179–97.
- Dawes, R. M. 1988. *Rational Choice in an Uncertain World*. San Diego: Harcourt Brace Jovanovich.
- De Quillacq, L. 1992. *The Power Brokers: An Insider's Guide to the French Financial Elite*. Dublin: Lafferty Publications.
- DiMaggio, P. 1992. "Nadel's Paradox Revisited: Relational and Cultural Aspects of Organizational Structure." Pp. 118–42 in *Organizations and Networks: Structure, Form, and Action*, edited by N. Nohria and R. G. Eccles. Boston: Harvard Business School Press.
- Durkheim, E. L. 1933. *Division of Labor in Society*. New York: MacMillan.
- Economist*. 1991. "Et Tu, Fabius." *Economist*, 14 December, 55–56.
- . 1992a. "After Mitterrand." *Economist*, 11 January, 64–65.
- . 1992b. "Slipping on Banana Skins." *Economist*, 4 April, 94–95.
- . 1992c. "A Right Mess." *Economist*, 16 May, 62–65.
- . 1992d. "The Big Squeeze." *Economist*, 17 June, 77–79.
- . 1992e. "The Big Squeeze." *Economist*, 17 October, 86.
- . 1993a. "End of the Road in France." *Economist*, 6 February, 54–55.
- . 1993b. "Patriot Games." *Economist*, 20 February, 72.
- . 1993c. "End of the Line in France." *Economist*, 8 May, 58.
- . 1993d. "Takeover." *Economist*, 13 November, 93–94.

- . 1994a. "Discredit Lyonnais." *Economist*, 2 April, 74.
- . 1994b. "An Affair of State." *Economist*, 16 July, 68–69.
- . 1995a. "Enarchy." *Economist*, 15 April, 47–48.
- . 1995b. "Jospin, the Surprise to Beat." *Economist*, 29 April, 57–58.
- . 1995c. "Franc Fortitude." *Economist*, 14 October, 92.
- . 1995d. "France Strikes Against Chirac." *Economist*, 14 October, 57–58.
- Emirbayer, Mustafa. 1997. "Manifesto for a Relational Sociology." *American Journal of Sociology* 103 (2): 281–317.
- Epstein, A. 1961 *The Network and Urban Social Organisation*. London: Her Majesty's Stationery Office.
- Feld, S. L. 1981. "The Focused Organization of Social Ties." *American Journal of Sociology* 86 (5): 1015–35.
- Festinger, L., S. Schachter, and K. Back. 1950. *Social Pressures in Informal Groups*. Stanford, Calif: Stanford University Press.
- Fienberg, S., M. Meyer, and S. Wasserman. 1985. "Statistical Analysis of Multiple Sociometric Relations." *Journal of the American Statistical Association* 80 (389): 51–67.
- Fienberg, S. E., and S. S. Wasserman. 1981. "Categorical Data Analysis of Single Sociometric Relations." Pp. 156–93 in *Analysis of Single Sociometric Relations*, edited by S. Leinhardt. San Francisco: Jossey-Bass.
- Frank, K. A. 1993. *Identifying Cohesive Subgroups*. Chicago: University of Chicago Press.
- . 1995. "Identifying Cohesive Subgroups." *Social Networks* 17:27–56.
- . 1996. "Mapping Interactions within and between Cohesive Subgroups." *Social Networks* 18:93–119.
- Frank, K. A., and J. Yasumoto. 1996. "Embedding Subgroups in the Sociogram: Linking Theory and Image." *Connections* 19 (1): 43–57.
- Frank, O. 1991. "Statistical Analyses of Changes in Networks." *Statistica Neerlandica* 45:283–93.
- Frank, O., and K. Nowicki. 1993. "Exploratory Statistical Analysis of Networks." *Annals of Discrete Mathematics* 55:349–66.
- Frank, O., and D. Strauss. 1986. "Markov Graphs." *Journal of the American Statistical Association* 81 (395): 832–42.
- Freeman, L. C. 1979. "Centrality in Social Networks Conceptual Clarification." *Social Networks* 1:215–39.
- . 1992a. "The Sociological Concept of 'Group': An Empirical Test of Two Models." *American Journal of Sociology* 98 (1): 152–66.
- . 1992b. "Filling in the Blanks: A Theory of Cognitive Categories and the Structure of Social Affiliation." *Social Psychology Quarterly* 55 (2): 118–27.
- Friedkin, Noah E. 1997. *A Structural Theory of Social Influence*. New York: Cambridge University Press.
- Goldstein, H. 1991. "Nonlinear Multilevel Models with an Application to Discrete Response Data." *Psychometrika* 78 (1): 45–51.
- Granovetter, M. 1973. "The Strength of Weak Ties: Network Theory Revisited." *American Journal of Sociology* 18 (4): 279–88.
- . 1985. "Economic Action and Social Structure: The Problem of Embeddedness." *American Journal of Sociology* 3:481–510.
- Hechter, Michael. 1983. "A Theory of Group Solidarity." Pp. 16–57 in *The Micro-foundations of Macrosociology*, edited by Michael Hechter. Philadelphia: Temple University Press.
- Hedeker, D., and R. Gibbons. 1996. "MIXOR: A Computer Program for Mixed-Effects Ordinal Probit and Logistic Regression Analysis." *Computer Methods and Programs in Biomedicine* 49:157–76.
- Heider, F. 1958. *The Psychology of Interpersonal Relations*. New York: Wiley.

- Higley, J., U. Hoffman-Lange, C. Kadushin, and G. Moore. 1991. "Elite Integration in Stable Democracies: A Reconsideration." *European Sociological Review* 7:35–53.
- Holland, P. W., and S. Leinhardt. 1981. "An Exponential Family of Probability Distributions for Directed Graphs." *Journal of American Statistical Association* 76 (373): 33–49.
- Homans, G. C. 1950. *The Human Group*. New York: Harcourt Brace.
- Hubert, L. J. 1987. *Assignment Methods in Combinatorial Data Analysis*. New York: Marcel Dekker.
- Kadushin, C. 1966. "The Friends and Supporters of Psychotherapy: On Social Circles in Urban Life." *American Sociological Review* 31:786–802.
- . 1995. "Friendship among the French Financial Elite." *American Sociological Review* 60:201–21.
- Laumann, E., and F. Pappi. 1976. *Networks of Collective Action: A Perspective on Community Influence*. New York: Academic Press.
- Lévi-Strauss, C. 1969. *The Elementary Structures of Kinship*. Boston: Beacon Press.
- Mariolis, P. 1982. "'Region' and 'Subgroup': Organizing Concepts in Social Network Analysis." *Social Networks* 4:305–28.
- Merton, R. K. 1957. *Social Theory and Social Structure*. Glencoe, Ill.: Free Press.
- Mudrack, P. E. 1989. "Defining Group Cohesiveness: A Legacy of Confusion?" *Small Group Behavior* 20 (1): 37–49.
- Nadel, S. F. 1957. *The Theory of Social Structure*. London: Cohen & West.
- Newcomb, T. M. 1961. *The Acquaintance Process*. New York: Holt, Rinehart & Winston.
- Newcomb, T. M., R. H. Turner, and P. E. Converse. 1965. *Social Psychology: The Study of Human Interaction*. New York: Holt, Rinehart & Winston.
- Olmsted, Michael S., and A. Paul Hare. 1978. *The Small Group*. New York: Random House.
- Portes, A. 1998. "Social Capital: Its Origins and Applications in Modern Sociology." *Annual Review of Sociology* 24:1–24.
- Portes, A., and J. Sensenbrenner. 1993. "Embeddedness and Immigration: Notes on the Social Determinants of Economic Action." *American Journal of Sociology* 98 (6): 1320–50.
- Putnam, R. 1993. "The Prosperous Community: Social Capital and Economic Growth." *Current* 356:4–6.
- Raub, Werner, and Jeroen Weesie. 1990. "Reputation and Efficiency in Social Interactions: An Example of Network Effects." *American Journal of Sociology* 96 (3): 626–54.
- Raudenbush, S. 1995. "Posterior Model Estimation for Hierarchical Generalized Linear Models with Application to Dichotomous and Count Data." Manuscript. College of Education, Michigan State University.
- Robinson, M. 1981. "The Identity of Human Social Groups." *Behavioral Science* 26: 114–29.
- Schlicht, K. 1984. "Cognitive Dissonance in Economics." Pp. 61–81 in *Normengesellschaftes Verhalten in Den Sozialwissenschaften*, edited by Duncker und Humbolt. Berlin.
- Selznick, P. 1961. "Foundations of the Theory of Organization." Pp. 18–31 in *Complex Organizations*, edited by A. Etzioni. New York: Holt, Rinehart & Winston.
- Shaw, M. E. 1981. *Group Dynamics: The Psychology of Small Group Behavior*. New York: McGraw Hill.
- Sherif, M., and C. W. Sherif. 1966. *Groups in Harmony and Tension: An Integration of Studies*. New York: Octagon Books.
- Simmel, G. 1955. *Conflict and the Web of Group Affiliations*. Glencoe, Ill.: Free Press.
- Simon, H. A. 1965. "The Architecture of Complexity." Pp. 63–76 in *General Systems*:

- Yearbook of the Society for General Systems*, vol. 10. Edited by L. von Bertalanffy and A. Rapaport. Ann Arbor, Mich.: Braun-Brumfeld.
- Smelser, N., and R. Swedberg. 1994. "The Sociological Perspective on the Economy." Pp. 3-26 in *The Handbook of Economic Sociology*, edited by N. Smelser and R. Swedberg. Princeton, N.J.: Princeton University Press.
- Strauss, D., and M. Ikeda. 1990. "Pseudolikelihood Estimation for Social Networks." *Journal of the American Statistical Association* 85 (409): 204-12.
- Uzzi, B. 1996. "The Sources and Consequences of Embeddedness for the Economic Performance of Organizations: The Network Effect." *American Sociological Review* 61: 674-98.
- von Furer-Haimendorf, C. 1938. "The *Morung* System of the Konyak Nagas." *Journal of the Royal Anthropological Institute* 68:349-78.
- Wasserman, Stanley, and K. Faust. 1994. *Social Networks Analysis: Methods and Applications*. New York: Cambridge University Press.
- Wasserman, S., and P. Pattison. 1996. "Logit Models and Logistic Regressions for Univariate and Bivariate Social Networks: I. An Introduction to Markov Graphs." *Psychometrika* 61 (3): 401-26.
- Weber, M. (1922) 1947. *The Theory of Social and Economic Organization*. New York: Free Press.
- Weick, K. 1993. "Sensemaking in Organizations: Small Structures with Large Consequences." Pp. 10-37 in *Social Psychology in Organizations: Advances in Theory and Research*, edited by J. K. Murnighan. Englewood Cliffs, N.J.: Prentice Hall.
- Wellman, B., and S. Wortley. 1990. "Different Strokes from Different Folks: Community Ties and Social Support." *American Journal of Sociology* 96 (3): 558-88.
- White, Harrison C. 1981. "Where Do Markets Come From?" *American Journal of Sociology* 87 517-47.
- Woolcock, M. 1998. "Social Capital and Economic Development: Towards a Theoretical Synthesis and Policy Framework." *Theory and Society* 27 (2): 151-208.
- Zafirovski, M., and B. Levine. 1997. "Economic Sociology Reformulated." *American Journal of Economic Sociology* 56 (3): 265-85.
- Zeger, S. L., K.-Y. Liang, and P. S. Albert. 1988. "Models for Longitudinal Data: A Generalized Estimating Equation Approach." *Biometrika* 44:1049-60.

Origins, Destinations, and Association in Occupational Mobility¹

Michael E. Sobel
University of Arizona

Mark P. Becker
University of Michigan

Susan M. Minick
Indiana University

Although the literature on occupational mobility distinguishes between structural and circulation (or exchange) mobility, previous methods for simultaneously considering these are inadequate. In this article, nonlinear models of association are combined with linear models for logits of marginal distributions, resulting in models with parameters directly interpretable in terms of association or structural mobility. Goodman's work on combining categories in mobility tables to handle the case where the marginal distributions are also of interest is extended. Extensions to sets of mobility tables are considered next. The article shows finally how to model cross-national variation in association and structural mobility.

INTRODUCTION

A long-standing distinction in the literature on intergenerational occupational mobility differentiates between mobility due to changes in the marginal distributions of origins and destinations, often called "structural mobility," and mobility that arises from the "openness" of the stratification system, often called "circulation mobility" ("pure mobility," "exchange mobility"). Until recently, in comparative stratification research, attention has focused primarily on the second type of mobility (which has been

¹ Portions of this work were completed while Sobel was on sabbatical at the University of Michigan. The authors are indebted to Yu Xie for stimulating discussion and to Yu Xie and several *AJS* reviewers for helpful comments on earlier drafts. Becker and Minick gratefully acknowledge support from the National Institutes of Health (CA53787). Correspondence should be directed to Michael Sobel, Department of Sociology, Social Sciences Building, Room 400, University of Arizona, Tucson, Arizona 85721

operationalized in a number of different ways), with most investigators treating structural mobility as a component of mobility that should be "partialled out" before two (or more) mobility tables are compared; many researchers using this approach have concluded that the patterns of association between fathers' and sons' occupations are cross-nationally "similar." At the same time, a number of investigators have pointed to the need to substantively address cross-national variation in structural mobility.

Sobel, Hout, and Duncan (1985) provide a framework for separating exchange and structural mobility that has proven useful. However, their approach requires the model of quasi symmetry (hereafter *QS*) to hold; under this model, the pattern of association is "symmetrical." In addition, structural mobility, understood as differences in origin and destination marginals, is not readily modeled within their framework because the marginal distributions are nonlinear functions of the model parameters. This paper proposes new methods for modeling structural mobility. Instead of modeling the expected frequencies in the mobility table using log-linear or log-nonlinear models of the type usually employed, we combine linear models for logits of the marginal distributions of origins and destinations with linear or nonlinear models for the pattern of association. Using this methodology, it is straightforward to simultaneously model cross-national and intertemporal variation in association between origins and destinations and cross-national and intertemporal variation in differences between origin and destination distributions.

To illustrate the utility of the proposed methods, three examples are presented. The purpose of the first is to introduce and compare our approach to modeling structural mobility with that in Sobel et al. (1985). Thus, we consider the Brazilian mobility data they used to demonstrate their approach. Further, as their approach requires *QS* to hold, in our analysis attention also focuses on models with symmetrical patterns of association. Our comparison demonstrates that even when *QS* holds, the framework used herein for modeling structural mobility is more "informative and less potentially misleading than using the structural mobility ratios" in Sobel et al. (1985). In addition to providing a parsimonious characterization of structural mobility in the Brazilian data, we also demonstrate the utility of using baseline-category logits for modeling structural mobility (as vs. cumulative or adjacent-category logits), and we model the pattern of association in the data using a model for ordinal data that is more parsimonious than the models typically used in mobility research.

The second example is a mobility table from Great Britain (Miller 1960). These data were used by Goodman (1981) to illustrate criteria for combining categories in cross-classifications of counts, generally, and, more specifically, in mobility tables. The pattern of mobility observed in a mobility table depends upon the occupational categories used. Two or more of the

categories may not be empirically distinct in some instances, in which case it might be desirable to combine them. Thus, approaches are needed for addressing the issue of category distinguishability or, equivalently, category collapsibility. Goodman's (1981) approach to addressing such issues focuses entirely on the pattern of association. We demonstrate the inadequacy of this approach when structural mobility is of interest, and we present a new criterion for evaluating when categories can be combined without changing either the pattern of association or the underlying pattern of structural mobility.

In our third example, we show how the methodology developed herein can be used to model sets of mobility tables, and we show how to model and test restrictions on the pattern of association and the form of structural mobility across the tables. The data are in a set of three mobility tables from Great Britain, Japan, and the United States, and they have been analyzed by Yamaguchi (1987), Xie (1992), and Lang and Eliason (1997). Our contributions complement these earlier contributions while providing new models and approaches for making cross-national comparisons of both association and structural mobility. We conclude with a discussion of other potentially useful applications of this methodology, and we briefly indicate the types of extensions that we shall consider in future work.

PRIOR RESEARCH

Before the advent of log-linear models in mobility research, two approaches to measuring circulation mobility were common. In the first, the discrepancy between the occupational distribution of the sons and the occupational distribution of the sons' fathers was operationalized as a lower bound on the observed mobility rate, in particular, as the smallest rate of mobility that must be observed in a mobility table by virtue of dissimilarity between the row and column margins of the table; circulation mobility, which measures the openness of the stratification system, was then equated with mobility in excess of this minimum. In the second approach (Rogoff 1953), social mobility ratios were computed. Using the model of statistical independence (perfect mobility) as a baseline for measuring "openness" for each cell in the mobility table, the ratio of the observed frequency to the frequency expected under perfect mobility was computed. A number of indices that combine features of both approaches were also proposed (Yasuda 1964; Boudon 1973). Problems with these approaches to the measurement of circulation mobility are well known (Duncan 1966; Hauser and Grusky 1988).

When researchers began using log-linear models over 20 years ago, the distinction between circulation and structural mobility was carried over

from previous work, with structural mobility reflected by marginal heterogeneity, as before, and circulation mobility equated with the pattern of association between fathers' and sons' occupations; the lack of correspondence between the notion of circulation implicit here and previous notions was pointed out by Sobel (1983). In early work (Hauser, Dickinson, Travis, and Koffel 1975; Hauser, Koffel, Travis, and Dickinson 1975), investigators found little evidence of cross-national or intertemporal variation in association, prompting them to propose that more attention be paid to variation in structural mobility. (Hauser and Featherman [1977] reiterate this suggestion.) However, about the same time, Featherman, Jones, and Hauser (1975) put forth the hypothesis that exchange mobility (understood as association) was spatiotemporally "similar." For at least 10 years thereafter, with the exception of Hope (1981, 1982), comparative researchers focused on this issue, to the neglect of the former proposal.

Subsequently, Sobel et al. (1985), following a lead from Hutchinson (1958), redefined exchange mobility as symmetric flows across the diagonal and showed how to separate exchange and structural mobility, under certain conditions. Specifically, let $i = 1, \dots, I$ index the respondent's origin, let $j = 1, \dots, I$ index the respondent's destination, let F_{ij} denote the expected frequency in cell (i, j) of the mobility table, and let

$$\log F_{ij} = \beta_i + \beta_j + \alpha_i + \delta_{ij}. \quad (1)$$

By definition, a pure exchange process holds if and only if $F_{ij} = F_{ji}$ for all i and j , that is, if and only if $\alpha_i = c$ for some constant c and $\delta_{ij} = \delta_{ji}$ for all i and j . Note that pure exchange implies marginal homogeneity, that is, the absence of structural mobility. Next, they consider QS, in which $\delta_{ij} = \delta_{ji}$ for all i and j , as above, but α_i is no longer a constant, implying that when QS holds, all marginal heterogeneity is attributable to the α_i parameters. Further, it may be shown that $F_{ij}/F_{ji} = c_i/c_j$ for all i and j , $i \neq j$, where c_i and c_j are unique positive constants, if and only if QS holds. Sobel et al. (1985) showed that when QS holds, $c_i/c_j = \exp(\alpha_i)/\exp(\alpha_j)$, that is, the structural mobility ratios $(F_{ij}/F_{ji}) = \exp(\alpha_i)/\exp(\alpha_j) = \exp(\alpha_i - \alpha_j)$ are uniquely identified, and asymmetry and structural mobility are equivalent.

Sobel (1988) extended this framework to higher way tables, including the case where a square table is observed in each of $k = 1, \dots, K$ settings:

$$\log F_{ijk} = \beta_{ik} + \beta_{jk} + \alpha_{ik} + \delta_{ijk}. \quad (2)$$

If, for every k , $\delta_{ijk} = \delta_{jik}$ for all i and j , QS holds in each table; under this condition, the α_{ik} account for marginal heterogeneity, and the ratios $(F_{ijk}/F_{jik}) = \exp(\alpha_{ik} - \alpha_{jk})$ are uniquely identified.

Luijkx (1994) criticized the approach in Sobel et al. (1985), arguing that insofar as a number of researchers (Hout 1989; Hout and Hauser 1992)

have incorrectly interpreted the α , parameters of (1) under *QS* as indicative of the contraction and expansion of occupations, marginal differences between origin and destination distributions should be modeled directly. Wong (1990) also interprets the α , parameters in this fashion. Thus, Luijkx (1994) proposed using generalized log-linear models (Becker and Balagtas 1993; Becker 1994; Lang and Agresti 1994; McCullagh and Nelder 1989), which combine log-linear models for the marginal distributions with log-linear models for the pattern of association, and he applied such a model to 41 educational heterogamy tables previously analyzed by Ultee and Luijkx (1990). For each table, $k = 1, \dots, 41$, he treats the row and column variables as ordered, fitting the pattern of association with the model of quasi-uniform association (Duncan 1979; Goodman 1979), and the marginal distributions with the linear shift model:

$$\log(F_{i,k}F_{i+1,k}/F_{\cdot,k}F_{i+1,\cdot}) = \alpha_i^*,$$

where the \cdot notation denotes summation over the relevant subscript; for example, $F_{\cdot,k} = \sum_{j=1}^J F_{j,k}$. He does not consider the case where one or both portions of the model are homogeneous across tables.

Subsequently, Lang and Eliason (1997) also used generalized log-linear models, arguing that Sobel et al.'s (1985) framework, because it requires *QS* to hold, is too restrictive. They considered ordered tables only and equated structural mobility with marginal heterogeneity, modeling the marginal distributions using linear models for the cumulative marginal logits and log-linear models for ordinal data, for example, the row-plus-column effects model with diagonal parameters (Goodman 1979, 1981), for the association. They did not explicitly consider models that use other kinds of logits for modeling the marginal distributions, nor did they consider nonlinear patterns of association, such as the so-called *RC* model and variations thereof.

Mathematically, the rationale for the use of generalized log-linear models stems from a result of Sinkhorn (1967), who shows that the marginal distributions and the nonredundant odds ratios of the two-way contingency table uniquely determine the cell probabilities in the table. Thus, combining a model for the marginal distributions with a model for the pattern of association implies a model for the joint distribution. However, we reject the substantive arguments put forth by Luijkx (1994) and Lang and Eliason (1997) for using generalized log-linear models in mobility research. Luijkx's argument is illogical: that subsequent researchers have misused the framework proposed by Sobel et al. (1985) does not imply the framework is problematic. Lang and Eliason's argument hinges on the lack of generality of the approach taken in Sobel et al. (1985). But this does not imply the utility of the approach they advocate, only that it can be applied to any table or set of tables; for example, one can always com-

pute the indices of structural and circulation mobility previously described.

Nevertheless, we believe that the approach based on combining models for marginal distributions with models for the pattern of association enables researchers to better analyze mobility tables. Thus, we provide a firmer rationale for this approach. We do so by comparing this approach with others, especially that of Sobel et al. (1985).

We also note that the methods used by Luijckx (1994) and Lang and Eliason (1997) are limited in that the models they use to describe the association in the mobility table are log linear. Thus, when models with row and column scores are used to describe the association, the scores must be known a priori. In many instances, this is not the case. For this reason, sociologists typically use log-nonlinear models of the type proposed by Goodman (1979) and extended by Clogg (1982) and Becker and Clogg (1989). Therefore, we use new models that combine nonlinear models for the pattern of association with linear models for the logits of the marginal distributions. With these models, it is straightforward to evaluate the consequences of combining two or more occupational categories. Finally, some authors (e.g., Goldthorpe 1980) have emphasized the importance of distinguishing between absolute and relative mobility rates. While our models directly parameterize association (relative mobility), we do not consider here models whose parameters are directly interpretable in terms of absolute mobility. Nevertheless, as our models imply a joint distribution, the fitted counts can be used to compute estimated rates of absolute mobility.

STRUCTURAL MOBILITY AND ASSOCIATION

Following many previous workers, we equate structural mobility with marginal heterogeneity between origin and destination distributions. Correspondingly, when we refer to shifts or flows from one occupational category to others in our empirical examples, we take these to be synonymous with relevant contrasts of marginal probabilities (or logits). We also assert that it is more illuminating to construct statistical models of mobility than to compute the types of indices used in earlier work, and we note that if such indices are desired, these may be obtained as a by-product of the modeling process.

To facilitate exposition, we begin with the case of a single mobility table. In the interest of generality, we consider the saturated model for the joint distribution, parameterized as in (1); note, however, that the results below apply also to unsaturated models, for example, to models with nonlinear patterns of association. As is well known, under the parameter-

ization (1), the marginal distributions are, in general, a function of all the model parameters:

$$\pi_i = F_i/F = \exp(\beta_i) \sum_{j=1}^I \exp(\beta_j + \alpha_j + \delta_{ij}) / \sum_{i=1}^I \left[\exp(\beta_i) \sum_{j=1}^I \exp(\beta_j + \alpha_j + \delta_{ij}) \right], \quad (3)$$

$$\pi_i = F_i/F = \exp(\beta_i + \alpha_i) \sum_{j=1}^I \exp(\beta_j + \delta_{ij}) / \sum_{i=1}^I \left[\exp(\beta_i + \alpha_i) \sum_{j=1}^I \exp(\beta_j + \delta_{ij}) \right]. \quad (4)$$

Thus, the parameters β_i , $i = 1, \dots, I$ and α_j , $j = 1, \dots, I$ do not refer to the marginal distributions unless special conditions hold. For example, under independence, the δ parameters equal zero, and it follows from (3) and (4) that $\pi_i = \exp(\beta_i) / \sum_{i=1}^I \exp(\beta_i)$ and $\pi_i = \exp(\beta_i + \alpha_i) / \sum_{i=1}^I \exp(\beta_i + \alpha_i)$.

Sobel et al. (1985) consider the special case of (1), where $\delta_{ij} = \delta_{jj}$ for all i and j ; here, the α_j parameters account for all marginal heterogeneity, and as previously noted the structural mobility ratios $F_{ij}/F_{jj} = \exp(\alpha_j) / \exp(\alpha_i)$ are invariant with respect to the manner in which the model is identified. Thus, Sobel et al. (1985) equate structural mobility with the α_j parameters. However, unlike several subsequent authors (Hout 1989; Hout and Hauser 1992; Wong 1990), they do not claim that $\alpha_j > 0$ ($\alpha_j < 0$) corresponds to expansion (contraction) of occupational category j . To see why this stronger interpretation is incorrect, note that (3) and (4) imply

$$\pi_j/\pi_i = \left[\sum_{i=1}^I \exp(\beta_i + \delta_{ij}) \right] / \left[\sum_{i=1}^I \exp(\beta_i + \alpha_i - \alpha_j + \delta_{ij}) \right],$$

that is, the expansion or contraction of occupational category j depends on all the model parameters, not just α_j . The correct interpretation of individual α_j parameters is in terms of an average departure, on a logarithmic scale, from symmetry for cells with a j subscript, since α_j is proportional to $\sum_{i \neq j} \log(F_{ij}/F_{jj})$ (Becker 1990b). Becker (1990b) notes that the minimum and maximum α parameters can be used to infer directions of marginal differences (i.e., contraction or expansion), but such inferences

are not valid more generally. That is, other than for the maximum and minimum α_j 's, it is indeed possible to have $\alpha_j > 0$ and yet $\pi_j/\pi_i < 1$.

A more straightforward approach is to model structural mobility (i.e., marginal heterogeneity) using logits of the marginal distributions, combining this with a model for the association, for example:

$$\log(F_i/F_I) = \gamma_i \quad (5)$$

where $i = 1, \dots, I - 1$,

$$\log(F_j/F_I) = \gamma_j + \lambda_j, \quad (6)$$

where $j = 1, \dots, I - 1$,

$$\log(F_{ij}F_{i+1,j+1}/F_{i,j+1}F_{i+1,j}) = \theta_{ij} \quad (7)$$

where $i = 1, \dots, I - 1$ and $j = 1, \dots, I - 1$. Though not parameterized as a log-linear model, this model is, due to Sinkhorn's (1967) result, equivalent to the saturated log-linear model (1), with the $I - 1$ baseline-category logits in (5) describing the origin distribution, the $I - 1$ baseline-category logits in (6) describing the destination distribution, and the $(I - 1)(J - 1)$ local log-odds ratios describing the pattern of association. The parameters λ_j are direct measures of structural mobility, $\exp(\lambda_j) = (F_j/F_I)(F_j/F_I)$, and they do not rely on a particular association pattern, such as QS , for a valid structural mobility interpretation. Recall that the α_j defined above in the framework of Sobel et al. (1985) are indirect measures of structural mobility since they are functions of the structural mobility ratios F_j/F_{j*} , which have a valid structural mobility interpretation only when QS holds.

When the categories are ordered, one might prefer to use a different set of logits: for example, adjacent-category logits, continuation-ratio logits, or cumulative logits. We discuss this further in the context of the examples. Similarly, any other set of nonredundant odds ratios could be used in (7). Unsaturated models are obtained by imposing restrictions on the parameters of (5)–(7). For example, Sobel et al.'s (1985) model of pure exchange (symmetry) is obtained when $\lambda_j = 0$, $j = 1, \dots, I - 1$, and $\theta_{ij} = \theta_{ji}$ for all i and j , while QS is obtained when only the latter set of restrictions is imposed. The model of marginal homogeneity is obtained when only the former set is imposed; Sobel et al. (1985) do not consider this model. As illustrated in our examples, many other models can be obtained by further modeling the θ_{ij} or the γ_i and λ_j .

If (5)–(7) are used as a basis for modeling the mobility table, the model parameters describe mobility in terms of the marginal distributions (from which structural mobility is derived) and association. We believe, without regard to whether or not QS holds, that parameterizing mobility in terms of structural mobility and association is more useful than the approach

in Sobel et al. (1985), which decomposes mobility into structural mobility and exchange mobility. In support of this view, we note, as demonstrated above, that structural mobility is not readily modeled within their framework. Second, the pattern of association in the mobility table measures the dependence of son's occupation on father's occupation, and in both the mobility literature and the status attainment literature, characterizing this dependence has been the primary concern for many years. In that vein, although Sobel et al. (1985) decompose mobility into structural mobility and exchange mobility, they neither model exchange mobility nor attach much substantive interest to this concept. In each of their examples, attention focuses on the α_j parameters and the contribution of association to exchange mobility but not on exchange mobility itself.

The foregoing remarks extend readily to the case where K mobility tables are observed, as in cross-national comparisons. Under the parameterization (2) for the joint distribution, the marginal distributions are a nonlinear function of the model parameters. If the association is symmetric in table k , the parameters α_{jk} , $j = 1, \dots, I$, account for marginal heterogeneity in this table; however, these parameters do not indicate whether or not there is expansion or contraction of occupational category j in table k . From (2), for $k \neq k^*$, when $\beta_{ik} + \beta_{jk} + \alpha_{jk} - (\beta_{ik}^* + \beta_{jk}^* + \alpha_{jk}^*) = c(k, k^*)$ and $\delta_{ijk} = \delta_{ijk}^*$ for all i and j , $\pi_{ik} = \pi_{ik}^*$ and $\pi_{jk} = \pi_{jk}^*$, $i = 1, \dots, I$. Thus, hypotheses that the origin and destination distributions are identical in two or more tables can be tested using the parameterization (2). However, once such hypotheses are rejected and interest turns to characterizing differences in marginal distributions across settings, this parameterization is no longer very useful. By way of contrast, we construct models of the form:

$$\log(F_{ik}/F_{Ik}) = \gamma_{ik} \quad (8)$$

where $i = 1, \dots, I - 1$ and $k = 1, \dots, K$,

$$\log(F_{jk}/F_{Ik}) = \gamma_{jk} + \lambda_{jkb} \quad (9)$$

where $j = 1, \dots, I - 1$ and $k = 1, \dots, K$,

$$\log(F_{ijk}F_{i+1,j+1,k}/F_{i,j+1,k}F_{i+1,j,k}) = \theta_{ijk} \quad (10)$$

where $i = 1, \dots, I - 1$, $j = 1, \dots, I - 1$, and $k = 1, \dots, K$.

By imposing restrictions in (8) and (9), differences in the origin and destination distributions both within and across tables can be tested and modeled. More parsimonious patterns of association are obtained by using linear and nonlinear models for the log-odds ratios in (10). An estimation algorithm for fitting models of this type with linear models for the log-odds ratios is described in Becker (1994). An extension of that algorithm

TABLE 1
OCCUPATIONAL MOBILITY IN SÃO PAULO BRAZIL, 1956

ORIGIN	DESTINATION						TOTAL
	1	2	3	4	5	6	
1. Professional and high administrative	33	12	10	3	0	0	58
2. Managerial and executive	14	25	16	3	2	0	60
3. Intermediate nonmanual	13	16	68	21	16	1	135
4. Lower nonmanual	6	30	39	74	61	7	217
5. Skilled manual	5	16	26	45	132	24	248
6 Semiskilled and unskilled manual	1	9	29	41	142	116	338
Total	72	108	188	187	353	148	1,056

for log-multiplicative association models is given in the appendix. Note also that although baseline-category logits and local log-odds ratios are modeled above, other logits and log-odds ratios can be used in the model construction process, as below.

EMPIRICAL EXAMPLES

Brazil

Table 1, from Hutchinson (1958), is the cross-classification of father's occupation by son's occupation for a sample of the labor force of São Paulo, Brazil, in 1956. The occupational categories are ordered with respect to prestige, with "1" indicating professional and high administrative, "2" indicating managerial and executive, "3," intermediate nonmanual, "4," lower nonmanual, "5," skilled manual, and "6," = semiskilled and unskilled manual. Sobel et al. (1985) and Sobel (1995) use log-linear models to describe the association in this table, showing that *QS*, as well as a number of nested models that take into account the ordering of the categories, fit the data well. Sobel et al. (1985) use the structural mobility ratios estimated using *QS* to make the following substantive claims: (a) structural mobility is predominantly upward, with the exception of mobility between categories 1 and 2 and categories 4 and 5, and (b) the "strongest force in the data favors upward mobility from semiskilled and unskilled manual origins" (Sobel et al. 1985, p. 367). In support of this last point, they present the estimated ratios $\exp(\hat{\alpha}_j)/\exp(\hat{\alpha}_6)$, for $j = 1, \dots, 5$, respectively. For future reference, the logarithms of these quantities are, respectively, 3.10, 3.15, 2.61, 1.68, 1.88.

To begin our analysis, we consider linear and nonlinear models of the

TABLE 2
ASSOCIATION MODELS FOR THE BRAZILIAN
DATA

Model	L^2	df	P -value
QS	14.16	10	.17
$RC + I$	9.04	10	.53
$(R + C) + I$	17.22	14	.24
$RC + 1$	12.93	15	.61
RC	52.27	16	.00
QUA	19.56	18	.36
$R = C + 1$	23.97	19	.20
$UA + 1$	26.56	23	.27

form (7) for describing the association in the Brazilian data. At this point, we do not impose restrictions on the parameters of (5) and (6), and thus the models we consider are equivalent to log-linear or log-nonlinear models for the joint distribution. Table 2 presents the likelihood ratio chi square (L^2), degrees of freedom (df), and P -values for a number of the association models considered. Three of the models do not feature symmetric association. Under $RC + I$ (Goodman 1979, 1981),

$$\theta_{ij} = \theta(u_{i+1} - u_i)(v_{j+1} - v_j) + (\tau_{ij} + \tau_{i+1,j+1} - \tau_{i,j+1} - \tau_{i+1,j}), \quad (11)$$

where the scores u_i and v_j are estimated in addition to the other model parameters, and $\tau_{rc} = 0$ if $r \neq c$, $r = 1, \dots, I$, $c = 1, \dots, I$. The fit of this model to the data is adequate, as is the fit of the model $RC + 1$, in which the nonzero τ_{rc} are constrained to have the common value τ ; further, $RC + I$ does not fit better at the .05 level using the conditional likelihood ratio test ($L^2[RC + 1] - L^2[RC + I] = 4.29$; $df = 4$; $P = 0.37$). However, model RC , in which the additional restriction $\tau = 0$ is imposed, does not fit the data adequately ($P = 0.00$).

We now compare $RC + 1$ to nested models featuring symmetric association. The model denoted $UA + 1$ is the special case of $RC + 1$ in which $u_i = v_i = i$; this model fits the data adequately, and $RC + 1$ does not fit better at the .05 level ($L^2[UA + 1] - L^2[RC + 1] = 13.63$; $df = 8$; $P = 0.09$). Thus, it appears that $UA + 1$ can be used to describe the association in the Brazilian table. Before reaching this conclusion, we also compare UA with three other symmetric association models. The model of quasi-uniform association (QUA) (Duncan 1979) replaces the single diagonal parameter of $UA + 1$ with I distinct diagonal parameters; QUA fails to fit

TABLE 3

COMPARISON OF ORIGIN AND DESTINATION DISTRIBUTIONS FOR THE BRAZILIAN DATA

EMPIRICAL LOGITS				
COMPARISON	Origin	Destination	DIFFERENCE	SE
Cumulative logits:				
1 vs. 2-6	-2.8456	-2.6146	.2304	.1332
1, 2 vs. 3-6	-2.0731	-1.5823	.4907	.0890
1-3 vs. 4-6	-1.1549	-.6257	.5293	.0657
1-4 vs. 5,6	-.2206	.1024	.3230	.0601
1-5 vs. 6	.7534	1.8140	1.0606	.0878
Adjacent-category logits:				
1 vs. 2	-.0339	-.4054	-.3715	.2079
2 vs. 3	-.8110	-.5543	.2568	.1765
3 vs. 4	-.4747	.0053	.4799	.1343
4 vs. 5	-.1335	-.6354	-.5018	.1138
5 vs. 6	-.3097	.8692	1.1789	.1111
Baseline-category logits:				
1 vs. 6	-1.7626	-.7205	1.0421	.1449
2 vs. 6	-1.7288	-.3151	1.4137	.1524
3 vs. 6	-.9178	.2393	1.1570	.1195
4 vs. 6	-.4432	.2339	.6771	.1171
5 vs. 6	-.3097	.8692	1.1789	.1111

the data better at the .05 level than $UA + 1$ ($L^2[UA + 1] - L^2[QUA] = 7$; $df = 5$; $P = 0.22$). Model $R = C + 1$ is the special case of $RC + 1$ in which the estimated row and column scores are homogeneous, that is, $u_i = v_i$, $i = 1, \dots, I$, and it is evident from inspection that $R = C + 1$ does not improve on $UA + 1$. Similarly, QS , the most general of the symmetric association models, does not fit better than $UA + 1$.

Having selected a model to describe the association, attention focuses on modeling the marginal distributions. Because the occupational categories are ordered, we may describe the marginal distributions and the difference between these using, for example, baseline-category logits (appropriate for both ordered and unordered data), cumulative logits (appropriate for ordered data), or adjacent-category logits (appropriate for ordered data). Our concern is to model the marginals using the logits that give the clearest picture of structural mobility in the table.

The upper panel of table 3 uses cumulative logits to characterize the distribution of origins and destinations and the difference between these logits to characterize the difference between the origin and destination

distributions. For $i = 1, \dots, I - 1$, the i th cumulative logit for the origin distribution is defined as $c_i = \log([F_1 + \dots + F_i]/[F_{i+1} + \dots + F_I])$; cumulative logits for the destination distribution are defined analogously. The second and third columns report, respectively, the sample estimates of the c_i and $c_{i,}$ while the fourth column is the estimate of the difference between the cumulative logits for sons and fathers, that is, the estimate of $c_i - c_{i,}$. Finally, in the fifth column, the standard error of this difference is reported. In the middle panel, adjacent-category logits $a_i = \log([F_i]/[F_{i+1}])$ and $a_{i,} = \log([F_i]/[F_{i+1}])$ are used to characterize the origin and destination distributions, respectively; thus, columns 2 and 3 are the sample estimates of a_i and $a_{i,}$ respectively, while the fourth column gives the estimated difference $a_i - a_{i,}$ and the fifth column reports the SE of the difference. Finally, in the lower panel, the calculations above are repeated using the baseline-category logits $\log([F_i]/[F_I])$ and $\log([F_i]/[F_I])$.

Inspection of the cumulative logits suggests a large shift out of semi-skilled and unskilled manual labor into the other categories (see the last row of entries in the top panel). Similarly, a shift out of this category into skilled manual labor is revealed using the adjacent-category logits. Using these logits, one sees, in addition, shifts into skilled manual labor and intermediate nonmanual labor from the lower nonmanual category; Sobel et al. (1985) also note this, using structural mobility ratios. The baseline-category logits highlight the large shift out of semiskilled and unskilled manual labor into all the other categories that Sobel et al. (1985) pointed to as the predominant feature of these data. However, using their structural mobility ratios (logarithms reported above), one would conclude the shift into professional and managerial work is greater than the shift into intermediate nonmanual work, which is in turn greater than the shift into lower nonmanual and skilled manual work. By way of contrast, using the marginal distributions directly, it appears that the shift out of semiskilled and unskilled manual labor into lower nonmanual labor is smaller than the shifts into the other categories, which appear to be approximately equal.

The preliminary analysis of the marginal distributions above suggests that the baseline-category logits and the adjacent-category logits present a clearer picture of the structural mobility in these data than do the cumulative logits. Further, when the prevailing shift in the marginals stems primarily from movement out of one category into all others, the baseline-category logits highlight this feature, while the adjacent-category logits do not. Thus, for these data (and many other mobility tables as well), the baseline-category logits seem to present the clearest picture of structural mobility. Attention now turns to formal modeling of these logits.

We combine the $UA + 1$ association model, in which the marginal dis-

tributions are unrestricted, with the restrictions on the marginal shifts suggested by the preceding discussion:

$$\log(F_j/F_i) = \begin{cases} \gamma_j + \lambda_j & j = 1, 2, 3, 5, \\ \gamma_i + \lambda_i, & \end{cases} \quad (12)$$

otherwise,

$$\log(F_{i,j+1}/F_{i+1,j}) = \begin{cases} 0 & \text{if } |i - j| > 1, \\ 0 - \tau & \text{if } |i - j| = 1, \\ 0 + 2\tau & \text{if } i = j, \end{cases} \quad (13)$$

where the γ parameters are defined in (5). For this model, $L^2 = 30.00$ on 26 df, with $P = 0.27$, indicating an adequate fit; further, as the model is nested under $UA + 1$, and the conditional likelihood ratio statistic has value 3.44 on 3 df, with $P = 0.33$, the restrictions on the marginal parameters are tenable. The maximum-likelihood estimates of the γ_i (SEs in parentheses) are $\hat{\gamma}_1 = -1.8505$ (0.1266), $\hat{\gamma}_2 = -1.5871$ (0.1098), $\hat{\gamma}_3 = -0.9402$ (0.0884), $\hat{\gamma}_4 = -0.4484$ (0.0858), $\hat{\gamma}_5 = -0.3199$ (0.0703), indicating the predominance of semiskilled and unskilled manual workers in the origin distribution. Further, the logits are increasing (with respect to category number), indicating an inverse relationship between occupational prevalence and occupational prestige within the origin distribution. Note that Sobel et al. (1985) do not report these results; in their analysis, the marginal parameters (e.g., α and β in [1]) do not relate directly to the marginal distributions. The estimate of λ is 1.1942 (0.0876), while the estimate of λ_i is 0.6837 (0.1159), implying that the odds of a son working in nonmanual labor as versus semiskilled and unskilled manual labor are $\exp(0.6837) = 1.9812$ the odds on a father working in nonmanual labor as opposed to semiskilled and unskilled manual labor. For any of the other categories, the sons' odds are $\exp(1.1942) = 3.3009$ times the size of the fathers' odds. The results point to a large shift out of semiskilled and unskilled manual labor, with equal relative shifts into all categories other than nonmanual labor. The remaining estimates are $\hat{\theta} = 0.3798$ (0.0290), $\hat{\tau} = 0.6361$ (0.0773), indicating a strong dependence of son's occupation on father's occupation, with strongest dependence in the associations involving the diagonal cells. Becker (1990b) shows that diagonal parameters, such as τ in (13) and those in QUA more generally, can have an interpretation within the context of the so-called mover-stayer model. The mover-stayer model posits diagonal cell entries to be a mixture of "stayers" (i.e., people who remain on the diagonal with probability one) and "movers" (i.e., people who have positive probability of change) that did not move. For the model fit here, with a single diagonal parameter, the interpretation is that

TABLE 4
OCCUPATIONAL MOBILITY IN GREAT BRITAIN, 1949

ORIGIN	DESTINATION								TOTAL
	1	2	3	4	5	6	7	8	
1. Professional and high administrative	50	19	26	8	7	11	6	2	129
2. Managerial and executive ..	16	40	34	18	11	20	8	3	150
3. Inspectional, supervisory, and other nonmanual (high grade)	12	35	65	66	35	88	23	21	345
4. Inspectional, supervisory, and other nonmanual (low grade)	11	20	58	110	40	183	64	32	518
5. Routine grades of nonmanual	2	8	12	23	25	46	28	12	156
6. Skilled manual	12	28	102	162	90	554	230	177	1,355
7. Semiskilled manual	0	6	19	40	21	158	143	71	458
8. Unskilled manual	0	3	14	32	15	126	91	106	387
Total	103	159	330	459	244	1,186	593	424	3,498

each category has the same percentage breakdown of diagonal cells into stayers and movers that did not move. Approximately 53% ($\exp[-\hat{\tau}] = 0.5294$) of diagonal cell entries (for each category) are estimated to be movers that did not move, and 47% are estimated to be stayers.

Great Britain

Table 4, from Miller (1960), is a cross-classification of father's occupation by son's occupation for Great Britain in 1949. The occupational categories are "1" for professional and high administrative, "2" for managerial and executive, "3" for inspectional, supervisory, and other nonmanual (high grade), "4" for inspectional, supervisory, and other nonmanual (low grade), "5" for routine grades of nonmanual, "6" for skilled manual, "7" for semi-skilled manual, and "8" for unskilled manual. These data have been analyzed for various purposes and in various forms (i.e., with certain categories combined) by Glass (1954), Duncan (1979), Goodman (1979, 1981), Hauser (1979), Clogg (1981), Sobel et al. (1985), and Becker (1990b), among others. In particular, Goodman (1981) used the data to illustrate criteria for determining whether certain categories in a cross-classification could be combined without changing the pattern of association.

Glass (1954) considered the 7×7 cross-classification, combining categories 5 and 6 of table 4. Goodman (1981), focusing on the pattern of associa-

tion in table 4, argued that combining those categories is inappropriate, but categories 4 and 5 may be combined. He proposed two types of criteria, "homogeneity" and "structural" (not to be confused with structural mobility), for assessing whether or not certain categories in a cross-classification can be combined given a one-to-one correspondence between the rows and columns. Suppose the adjacent categories i and $i + 1$ are to be considered as candidates for being combined. A homogeneity condition is satisfied if quasi independence holds in the subtable formed by blanking out diagonal cells, (i, i) and $(i + 1, i + 1)$, and deleting cells that are not in rows or columns for categories i or $i + 1$. A structural criterion assesses if and how the association model of the full cross-classification is preserved when two categories are combined. Given a model for the full table in which the association model specifies quasi independence between categories i and $i + 1$, the likelihood ratio statistic for the model is partitioned into two components, one for testing independence in the subtable and the second testing the fit of a model that maintains the association model specified in the full table to the collapsed table (formed by combining categories i and $i + 1$). Thus, one can assess both homogeneity and structural criteria (based on the association) for the full table by examining the corresponding components of the likelihood ratio statistic. Applying these criteria, Goodman concluded that it is appropriate to combine categories 4 and 5 (see below) but that the homogeneity criterion is not satisfied for categories 5 and 6; hence, combining those categories provides a view of the pattern of association that is not congruent with the full cross-classification reported here.

Given our focus on structural mobility, it is inadequate to examine only the pattern of association in assessing the loss of information resulting from combining occupational categories. We propose an additional criterion to augment Goodman's (1981) criteria for combining categories by applying the same basic principles to the marginal distributions, and we note that his structural criterion is readily modified to incorporate the fact that the models we consider here include specifications both for association and for structural mobility. We illustrate the new structural criterion in the context of combining adjacent categories and models for adjacent-category logits. For some results on cumulative logits and continuation-ratio logits, see Minick (1997), who also proves an analogous result for combining arbitrary pairs of categories using adjacent-category logits. Note that combining an arbitrary pair of categories is inappropriate for cumulative logits and continuation-ratio logits since these logits are meaningful only when the response categories are strictly ordered.

Consider two categories i and $i + 1$. Using the adjacent-category logits, structural mobility is preserved when these categories are combined if

$$\begin{aligned} \log(\pi_i/\pi_{i-1}) - \log(\pi_i/\pi_{i-1}) &= \log[(\pi_i + \pi_{i+1})/\pi_{i-1}] \\ &- \log[(\pi_i + \pi_{i+1})/\pi_{i-1}] \end{aligned} \quad (14)$$

and

$$\begin{aligned} \log(\pi_{i+1}/\pi_{i+1}) - \log(\pi_{i+1}/\pi_{i+1}) &= \log[\pi_{i+1}/(\pi_i + \pi_{i+1})] \\ &- \log[\pi_{i+1}/(\pi_i + \pi_{i+1})] \end{aligned} \quad (15)$$

hold. Note that no other differences in origin and destination adjacent-category logits make use of the combined categories. Next, note that (14) holds if and only if

$$\frac{\pi_i}{\pi_i} = \frac{\pi_i + \pi_{i+1}}{\pi_i + \pi_{i+1}}. \quad (16)$$

Similarly, (15) holds if and only if

$$\frac{\pi_{i+1}}{\pi_{i+1}} = \frac{\pi_i + \pi_{i+1}}{\pi_i + \pi_{i+1}}. \quad (17)$$

From (16) and (17), it follows that the model for structural mobility is preserved if and only if

$$\frac{\pi_{i+1}}{\pi_i} = \frac{\pi_{i+1}}{\pi_i} \quad (18)$$

holds. The analogous criterion for an arbitrary pair of categories i and i' is $\pi_i/\pi_{i'} = \pi_i/\pi_{i'}$ (Minick 1997).

Given a model for the full table in which the association model specifies quasi independence between categories i and $i + 1$ and the model for the univariate marginals satisfies (18), the likelihood-ratio statistic for the model of the full table can be partitioned into three components corresponding to: (a) quasi independence in the subtable, (b) equality of the ratios of the marginal responses of the two categories (i.e., [18]), and (c) the fit of the model to the collapsed table that maintains the association and univariate marginal model specified for the full table. Thus, the homogeneity and structural criteria for the full table, now based on both association and univariate marginal models, can be examined using the corresponding components of the likelihood ratio statistic.

To begin our analysis, we consider the $RC + I$ association model (see [11] above and Goodman [1981]). This model provides a reasonable fit to the data ($L^2[RC + I] = 29.15$; $df = 28$; $P = 0.40$). The model equating the row and column scores—that is, $\mu_i = v_i$ for all i , denoted here as $R = C + I$ —is also tenable ($L^2[R = C + I] = 32.56$; $df = 34$; $P = 0.54$); in this model, the association is symmetric, and hence QS holds. Further,

comparison with the $RC + I$ association model also suggests that symmetric association is tenable; that is, $L^2(R = C + I) - L^2(RC + I) = 3.41$; $df = 6$, $P = 0.76$. However, QUA is not tenable: $L^2(QUA) = 58.44$; $df = 40$; $P = 0.03$, and $L^2(QUA) - L^2(R = C + I) = 25.88$; $df = 6$; $P = 0.00$. Thus, attention is focused on the nonlinear association model $R = C + I$.

The estimated scores from the $R = C + I$ model are $\mu_1 = 1$, $\hat{\mu}_2 = 1.59$, $\hat{\mu}_3 = 3.21$, $\hat{\mu}_4 = 4.79$, $\hat{\mu}_5 = 4.83$, $\hat{\mu}_6 = 6.22$, $\hat{\mu}_7 = 7.34$, and $\mu_8 = 8$, where the identification restrictions $\mu_1 = 1$ and $\mu_8 = 8$ are used to facilitate comparison to the uniformly spaced scores $\mu_i = i$. Alternative identifying restrictions, such as $\sum_i \mu_i = 0$ and $\sum_i \mu_i^2 = 1$, may be imposed by applying the appropriate location and scale adjustments to the scores reported here (see, e.g., Becker and Clogg [1989] and Becker [1990a]). The natural question to ask at this point is whether or not categories 4 and 5 are distinguishable from one another and whether or not the pattern of association or the structural mobility is altered by collapsing these categories to form a 7×7 table. Starting with the association, the likelihood ratio statistic for the model that specifies quasi independence for categories 4 and 5 and $R = C + I$ for the remainder of the table is partitioned into two components, one testing quasi independence in the subtable for categories 4 and 5, with diagonal deleted ($L^2 = 7.87$; $df = 11$; $P = 0.65$), and one testing $R = C + I$ in the collapsed 7×7 table ($L^2 = 32.56$; $df = 34$; $P = 0.54$). As in Goodman (1981), both the homogeneity and structural criteria are satisfied, leading to the conclusion that combining categories 4 and 5 yields no significant loss of information about the association in this table.

To consider both structural mobility and association in table 4, we consider models where the likelihood ratio statistic is partitioned into the following components: (a) quasi independence for categories 4 and 5, (b) the $R = C + I$ association model for the remainder of the table, and (c) equality of logits $\log(\pi_4/\pi_5)$ and $\log(\pi_4/\pi_3)$. The first two components were computed in the analysis of association above. The latter component yields $L^2 = 22.47$; $df = 1$; $P = 0.00$, and hence it does not support combining categories 4 and 5. Note that an asymptotically equivalent test may be obtained by dividing the "4 vs. 5" adjacent category logit difference by its estimated SE (middle panel of table 5) and then squaring; i.e., $(-0.5683/0.1168)^2 = 23.67$; $df = 1$; $P = 0.00$. The inference is that combining occupational categories 4 and 5 in table 4 results in a significant loss of information about differences between the origin and destination distributions, which in turn distorts the view of structural mobility.

Inspecting table 5, it is clear from the adjacent category logit differences that there is not a loss of information about structural mobility if categories 2 and 3 are combined $[(0.1026/0.1304)]^2 = 0.62$; $df = 1$; $P = 0.45$) or if categories 3 and 4 are combined $[(0.0766/0.0970)]^2 = 0.62$; $df = 1$; $P =$

TABLE 5

COMPARISON OF ORIGIN AND DESTINATION DISTRIBUTIONS FOR THE BRITISH DATA

	EMPIRICAL LOGITS			
COMPARISON	Origin	Destination	DIFFERENCE	SE
Cumulative logits				
1 vs. 2-8	-3.2623	-3.4966	-.2328	.1030
1, 2 vs 3-8	-2.4453	-2.5133	-.0680	.0688
1-3 vs 4-8	-1.5274	-1.5911	-.0637	.0498
1-4 vs 5-8	-.7242	-.8451	-.1209	.0423
1-5 vs. 6-8	-.5276	-.5314	-.0037	.0405
1-6 vs. 7,8	1.1441	.8918	-.2523	.0475
1-7 vs. 8 ..	2.0843	1.9810	-.1033	.0683
Adjacent-category logits				
1 vs. 2	-.1508	-.4342	-.2834	.1523
2 vs. 3	-.8329	-.7302	.1026	.1304
3 vs. 4	-.4065	-.3299	.0766	.0970
4 vs. 5	1.2001	.6319	-.5683	.1168
5 vs. 6	-2.1620	-1.5813	.5806	.1057
6 vs. 7	1.0847	.6931	-.3915	.0698
7 vs 81685	.3355	.1670	.0892
Baseline-category logits				
1 vs. 8	-1.0987	-1.4151	-.3164	.1169
2 vs. 8	-.9478	-.9808	-.0330	.1158
3 vs. 8	-.1148	-.2506	-.1358	.0941
4 vs. 82915	.0793	-.2123	.0870
5 vs. 8 ..	-.9086	-.5525	.3561	.1166
6 vs. 8	1.2531	1.0286	-.2245	.0756
7 vs. 81685	.3355	.1670	.0892

0.45). Combining either pair of categories (or all three categories) does, however, change the association structure of the table.

We now turn to modeling the marginal distributions in terms of adjacent-category logits. Let

$$\log(\pi_j/\pi_{j+1}) - \log(\pi_j/\pi_{j+1}) = \lambda_p$$

where $j = 1, \dots, 7$. The first reduced model (B_1) we consider for the marginal distributions sets $\lambda_1 = \lambda_2 = \lambda_3 = 0$. This model specifies no difference between origins and destinations in terms of the relative odds of the first four occupational categories, and it provides an excellent fit ($L^2[B_1] = 36.69$; $df = 37$; $P = 0.48$). Recall from above that $L^2 = 32.56$; $df = 34$, for the full $R = C + I$ model (hereafter B_0). It is evident from $L^2(B_1) - L^2(B_0) = 36.69 - 32.56 = 4.13$; $df = 37 - 34 = 3$; $P = 0.25$

that the simpler model fits just as well as the more general model. Further restricting the model so that the relative odds of categories 7 and 8 are the same in both distributions (i.e., $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_7 = 0$) results in model B_1 ($L^2[B_1] = 40.21$; $df = 38$; $P = 0.37$). Comparing the two reduced models ($L^2[B_1] - L^2[B_2] = 40.21 - 36.69 = 3.52$; $df = 1$; $P = 0.06$), there appears to be weak support for retaining the restriction $\lambda_7 = 0$. Thus, we arrive at a model specifying that structural mobility in Britain can be attributed to decreases from origins to destinations in the relative odds of categories 4 versus 5, 6 versus 7, and an increase in the relative odds of 5 versus 6; for example, see estimates in table 5. Note that while occupation category 6 is modal in both the origin and destination distributions, the relative dominance of that category is less in the destination distribution.

GREAT BRITAIN, JAPAN, AND THE UNITED STATES

Table 6, from Yamaguchi (1987), is a three-way cross-classification of father's occupation by son's occupation by nation. The occupational categories are "1" for upper nonmanual (professionals, managers, and officials), "2" for lower nonmanual (proprietors, sales workers, and clerical workers), "3" for upper manual (skilled workers), "4" for lower manual (semiskilled and unskilled nonfarm workers), and "5" for farm (farm workers and laborers). The data for the United States are from the 1973 Occupational Changes in a Generation Survey, the British data are from the 1972 Oxford Social Mobility Survey, and the Japanese data are from the 1975 Social Stratification and Mobility Survey. Comparisons of the American and British data are given in Kerckhoff, Campbell, and Winfield-Laired (1985), and all three tables are compared in Yamaguchi (1987), Xie (1992), and Lang and Eliason (1997).

Using log-linear models, Yamaguchi (1987) extends the framework in Sobel et al. (1985) to the case where QS holds in each of several tables. However, he cannot use the models to address structural mobility for the data in table 6 because QS does not hold in each of the three tables. Lang and Eliason (1997) use linear models to describe the pattern of association in the three tables, and they use cumulative logit models to compare origin and destination distributions. Xie (1992) uses log-multiplicative association models (Goodman 1979, Clogg 1982) to compare the patterns and strengths of association in the three tables, essentially ignoring the issue of structural mobility.

Our analysis complements those above, but it is different in several important respects. First, although the log-linear $R + C$ models of association in Yamaguchi (1987) and Lang and Eliason (1997) are easily accommodated within our framework, we focus on log-multiplicative models of association because these are more substantively interpretable. See Xie

TABLE 6

OCCUPATIONAL MOBILITY IN THE UNITED STATES (1973), GREAT BRITAIN (1972), AND JAPAN (1975)

ORIGIN	DESTINATION					TOTAL
	1	2	3	4	5	
United States:						
1. Upper nonmanual	1,275	364	274	272	17	2,202
2. Lower nonmanual	1,055	597	394	443	31	2,520
3. Upper manual	1,043	587	1,045	951	47	3,673
4. Lower manual	1,159	791	1,323	2,046	52	5,371
5. Farm	666	496	1,031	1,632	646	4,471
Total	5,198	2,835	4,067	5,344	793	18,237
Great Britain:						
1. Upper nonmanual	474	129	87	124	11	825
2. Lower nonmanual	300	218	171	220	8	917
3. Upper manual	438	254	669	703	16	2,080
4. Lower manual	601	388	932	1,789	37	3,747
5. Farm	76	56	125	295	191	743
Total	1,889	1,045	1,984	3,131	263	8,312
Japan:						
1. Upper nonmanual	127	101	24	30	12	294
2. Lower nonmanual	86	207	64	61	13	431
3. Upper manual	43	73	122	60	13	311
4. Lower manual	35	51	62	66	11	225
5. Farm	109	206	184	253	325	1,027
Total	400	638	456	470	374	2,338

(1992) for similar remarks pertaining directly to the data in table 6. However, the association models we consider are more general than those considered by Xie (1992), and we find that one of the more general models herein fits the data better than Xie's (1992) preferred model. Second, our analysis suggests that the occupational categories do not have the same ordering in both the origin and destination distributions in all three countries, so using cumulative logits to model structural mobility, as in Lang and Eliason (1997), is debatable. (The ordering of the scores is dependent upon the association model used; see, e.g., n. 1 in Lang and Eliason [1997].) Even were this not the case, baseline-category logits are clearly more informative and useful for modeling cross-national variation in structural mobility than the cumulative logits, as illustrated subsequently. Results using adjacent-category logits are also included, for completeness.

The general log-multiplicative *RC* association model for a set of two-way tables is

$$\theta_{ijk} = \theta_k(\mu_{i+1,k} - \mu_{ik})(v_{j+1,k} - v_{jk}),$$

where the μ_{ik} and v_{jk} are origin and destination scores, respectively, that are to be estimated from the data, and the θ_k parameters summarize the strength of association. See Clogg (1982), Becker and Clogg (1989), Becker (1990a), and Xie (1992) for discussion of identifying restrictions for these scores. Xie (1992) considered the special case where the origin and destination scores are restricted to be the same for all K tables (i.e., $\mu_{ik} = \mu_i$, $v_{jk} = v_j$), thus implying a homogeneous pattern of association with variable strengths across tables. We use the notation $(RC)_k$ for the purpose of summarizing assumptions about the score parameters in our models and $(SA)_k$ for summarizing assumptions about the strength of association θ_k . Consistent with notation adopted by Xie (1992), models where parameters are constrained to be homogeneous across nations are indicated by setting $k = 0$, and models where parameters are allowed to vary across nations are indicated by setting $k = x$. For example, a model with homogeneous scores and heterogeneous strength of association includes $(RC)_0$ and $(SA)_x$, respectively. All models considered here assume that the diagonal cells are treated separately, and we use the notation $(I)_x$ to indicate such. The four models of association that we consider are

$$A_1: (RC)_0 + (SA)_0 + (I)_x,$$

$$A_2: (RC)_0 + (SA)_x + (I)_x,$$

$$A_3: (RC)_x + (SA)_0 + (I)_x$$

and

$$A_4: (RC)_x + (SA)_x + (I)_x.$$

The first two models were considered by Xie (1992), and the latter two appear to be new in the mobility literature. Goodness-of-fit summaries for these models are reported in table 7, along with summaries of certain model comparisons.

It is evident from the results in table 7 that all four models fit the data reasonably well. However, in comparing models where the strength of association is homogeneous by nation, model A_3 is preferable to model A_1 ($L^2[A_1] - L^2[A_3] = 24.30$; $df = 12$; $P = 0.0185$), and we therefore reject the assumption of homogeneous scores employed by Xie (1992). In addition, there is little evidence to recommend model A_4 over model A_3 , so subsequent analysis focuses on A_3 . The origin and destination scores estimated under model A_3 are reported in table 8. Note that the identifying restrictions used to arrive at the estimates in table 8 are $\mu_{1k} = 1$, $\mu_{5k} = 5$,

TABLE 7
ASSOCIATION MODELS APPLIED TO TABLE 6

Description	Model	L^2	df	P-value
Homogeneous row and column scores and $\theta_i = \theta$	$A_1: (RC)_s + (SA)_s + (I)_s$	37.72	26	.06
Homogeneous row and column scores with unequal θ_i	$A_2: (RC)_s + (SA)_s + (I)_s$	32.12	24	.12
Heterogeneous row and column scores and $\theta_i = \theta$	$A_3: (RC)_s + (SA)_s + (I)_s$	13.42	14	.49
Heterogeneous row and column scores with heterogeneous θ_i	$A_4: (RC)_s + (SA)_s + (I)_s$	13.29	12	.35
Test assumption of homogeneous θ_i given $(RC)_s, (I)_s$	A_1 vs. A_2	4.60	2	.10
Test assumption of homogeneous θ_i given $(RC)_s, (I)_s$	A_2 vs. A_4	.13	2	.94
Test assumption of homogeneous scores given $(SA)_s, (I)_s$	A_1 vs. A_3	23.3	12	.03
Test assumption of homogeneous scores given $(SA)_s, (I)_s$	A_3 vs. A_4	18.83	12	.09

NOTE.—RC = row and column effects model; SA = strength of association as measured by θ_i ; I = I diagonal parameters; ()_s = homogeneous across tables; ()_h = heterogeneous across tables

$v_{14} = 1$, and $v_{54} = 5$. Xie (1992) reported an ordering violation for category 5 of the destination scores, and in estimating heterogeneous scores we find further evidence that the violation holds for all three nations. In addition, we find a violation for origin categories 4 and 5 in Japan, with near equality in the scores for these two occupational categories. The smallest difference between adjacent origin scores is indeed between categories 4 and 5 for both Britain and Japan, but in the United States the smallest difference is between categories 1 and 2. There appears to be polarization of the origin scores away from the middle category for all three nations, and the national differences in origin scores are largely in terms of the comparisons between categories 1 and 2 and between categories 4 and 5. The destination scores show a somewhat different pattern. The distance between categories 1 and 2 is the greatest in the United States and smallest in Japan. Reordering the occupational categories to place category 5 between categories 2 and 3, based on the estimates, the largest difference in destination scores is between categories 3 and 5 in the United States, between categories 3 and 4 for Britain, with approximate equality for both categories 2 and 5, and for 3 and 5 in Japan.

Given a model for the association, A_3 , attention now focuses on the marginal distributions and structural mobility. Tables 9, 10, and 11 report the empirical cumulative, adjacent-category, and baseline-category logits,

TABLE 8
ESTIMATED SCORES FOR ASSOCIATION IN MODEL A₁ OF TABLE 7

	ORIGIN SCORES					DESTINATION SCORES				
United States	1	1.46 (.33)	3.03 (.18)	4.22 (.15)	5	1	3.89 (.99)	9.26 (2.26)	11.48 (2.87)	5
Great Britain	1	1.99 (.40)	3.42 (.23)	4.88 (.36)	5	1	2.65 (1.26)	8.88 (2.30)	12.14 (3.14)	5
Japan	1	2.44 (.76)	3.91 (.47)	5.56 (.62)	5	1	1.80 (1.92)	8.19 (2.27)	9.76 (2.65)	5

NOTE.—Estimated SEs are in parentheses

TABLE 9

CUMULATIVE LOGITS FOR UNITED STATES, GREAT BRITAIN, AND JAPAN

EMPIRICAL LOGITS				
COMPARISON	Origin	Destination	DIFFERENCE	SE
United States:				
1 vs. 2-5	-1.9856	- 9195	1.0657	.0246
1, 2 vs. 3-5	-1.0515	-.2393	.8123	.0188
1-3 vs. 4-5	-.1590	.6788	.8379	.0187
1-4 vs. 5	1 1246	3 0909	1.9663	.0355
Great Britain:				
1 vs. 2-5	-2.2055	-1.2238	.9817	.0388
1, 2 vs. 3-5	-1 3276	-.6059	.7215	.0294
1-3 vs. 4-5	-.1611	.3709	.5320	.0274
1-4 vs. 5	2.3211	3.4212	1.1000	.0588
Japan:				
1 vs. 2-5	-1.9393	-1.5779	.3611	.0715
1, 2 vs. 3-5	-.7996	-.2250	.5746	.0485
1-3 vs. 4-5	-.2285	.5710	.7996	.0491
1-4 vs. 51577	1 6585	1.5008	.0568

TABLE 10

ADJACENT-CATEGORY LOGITS FOR THE UNITED STATES, GREAT BRITAIN, AND JAPAN

EMPIRICAL LOGITS				
COMPARISON	Origin	Destination	DIFFERENCE	SE
United States:				
1 vs. 2	-.1349	.6062	.7411	.0358
2 vs. 3	-.3767	-.3608	.0159	.0340
3 vs. 4	-.3799	-.2731	.1069	.0288
4 vs. 51834	1.9079	1.7245	.0389
Great Britain:				
1 vs. 2	-.1057	.5921	.6978	.0581
2 vs. 3	-.8189	-.6411	.1779	.0516
3 vs. 4	-.5886	-.4562	.1324	.0375
4 vs. 5	1.6180	2.4770	.8589	.0616
Japan:				
1 vs. 2	- 3826	-.4668	-.0843	.0905
2 vs. 33263	.3358	.0095	.0860
3 vs. 43237	-.0303	-.3539	.1051
4 vs. 5	-1.5659	.2285	1.7943	.0926

TABLE 11

BASELINE-CATEGORY LOGITS FOR THE UNITED STATES, GREAT BRITAIN, AND JAPAN

EMPIRICAL LOGITS				
COMPARISON	Origin	Destination	Difference	SE
United States:				
1 vs. 5	-7083	1.8802	2.5885	.0403
2 vs. 5	-5734	1.2740	1.8473	.0426
3 vs. 5	-1966	1.6348	1.8314	.0405
4 vs. 5	1834	1.9079	1.7245	.0389
Great Britain:				
1 vs. 5	.1047	1.9716	1.8670	.0671
2 vs. 5	.2104	1.3796	1.1692	.0707
3 vs. 5	1.0294	2.0207	.9913	.0640
4 vs. 5	1.6180	2.4770	.8589	.0616
Japan:				
1 vs. 5	-1.2983	.0672	1.3656	.0802
2 vs. 5	-.9158	.5341	1.4499	.0718
3 vs. 5	-1.2420	.1983	1.4404	.0814
4 vs. 5	-1.5659	.2285	1.7943	.0926

respectively, for each of the six marginal distributions. The fifth category, farming, is taken as the baseline-category in table 11. Note that Lang and Eliason (1997) model cumulative logits in their analysis of these data, that it is necessary to assume the categories are appropriately ordered in the table when using these logits, and that our analysis suggests it is doubtful the destination distribution is appropriately ordered. Here, we demonstrate how parsimonious baseline-category logit models of structural mobility can be constructed.

Lang and Eliason (1997) noted that the origin-destination differences indicate a greater shift toward the higher status categories in the United States. That is, all four of the cumulative logit differences for the United States in table 9 are greater than the corresponding differences for Britain and Japan. In comparing differences for Britain and Japan, the picture is slightly more complicated. The shift out of farming is greater for Japan (i.e., $\hat{\alpha}_{4(Japan)} - \hat{\alpha}_{1(Japan)} = 1.50 > 1.10 = \hat{\alpha}_{4(GB)} - \hat{\alpha}_{1(GB)}$), but net of farming there is more upward mobility in Britain. However, it is difficult to subscribe to this interpretation as upward mobility since the estimated scores from the association model raise the possibility that the ordering of the occupation categories differs for origins and destinations.

We now analyze these data using models specified in terms of baseline-category logits. The farm category is taken to be the baseline category

since flows out of this occupational category were dramatic during the period studied. The logit differences are uniformly larger for the United States than for Britain. The logit differences are greater for the United States than for Japan for all categories except lower manual, where the size of the discrepancy is trivial (1.72 versus 1.79). Note that other differences in relative log odds can be compared by taking appropriate differences of the logit differences reported in table 11. For example, in the United States, the log odds of 2 versus 3, lower nonmanual versus upper manual, equals $1.85 - 1.83 = 0.02$, thus indicating that the odds of category 2 versus category 3 are virtually identical in the origin and destination distributions. A similar phenomena is observed for Japan ($1.45 - 1.44 = 0.01$), while for Britain there is evidence of a slight increase in the destination distribution ($1.17 - 0.99 = 0.18$).

Another important observation to make from table 11 is that the pattern of origin-destination differences is relatively similar in the United States and Britain. That is, the largest change relative to farming is in the first category, and the differences then decrease in each subsequent category. However, as can be inferred from the comparison of categories 2 and 3 above, it is only the pattern and not the magnitudes that are the same in the United States and Britain. The picture for Japan is markedly different. There the greatest change in log odds is for category 4, and the differences for the other three categories are relatively similar in magnitude. We now consider a number of restricted models for the marginal distributions to illustrate how to formulate and test models of cross-national differences in occupational mobility.

The reduced models considered here are by no means exhaustive; they are chosen to illustrate how to formulate models for investigating the types of phenomena reported above for these data. We start with (8) and an additive model for differences between origins and destinations:

$$\log(F_{jk}/F_{jk}) = \gamma_j + \lambda_j + \delta_k \quad (19)$$

where $j = 1, \dots, 4$ and $k = 1, 2, 3$. That is, the $(I - 1) \times K$ parameters λ_{jk} in (9) are replaced by an additive function of $I - 1$ parameters λ_j , corresponding to logits, and K parameters δ_k , corresponding to nations. This model, hereafter M_1 , allows for differences between origin and destination that depend additively on both the logit number (1, 2, 3, 4) and nation. Note that an identifying restriction is required for δ_k . We use $\delta_1 = 0$, unless noted otherwise. For M_1 , $L^2(M_1) = 205.21$; $df = 20$; $P = 0.00$. Note that adding the interaction term λ_{jk} to M_1 yields A_1 , our preferred association model; under A_1 , $L^2(A_1) = 13.42$; $df = 14$; $P = 0.49$.

We now consider three different marginal models that exploit the aforementioned similarities between the United States and Britain. Under M_1 ,

both nations follow the same pattern but with a difference in magnitudes that is the same across all four logits:

$$\log(F_{jk}/F_{sk}) = \gamma_{jk} + \lambda_j + \delta_k,$$

where $j = 1, \dots, 4$ and $k = 1, 2$,

$$\log(F_{js}/F_{ss}) = \gamma_{js} + \lambda_{js},$$

where $j = 1, \dots, 4$. This model fits the data better than M_1 ($L^2[M_1] = 31.53$; $df = 17$; $P = 0.02$) but is still inadequate. Relaxing the assumption on the logit for category 4 yields model M_3 :

$$\log(F_{jk}/F_{sk}) = \gamma_{jk} + \lambda_j + \delta_k,$$

where $j = 1, 2, 3$ and $k = 1, 2$,

$$\log(F_{4k}/F_{sk}) = \gamma_{4k} + \lambda_{4k},$$

where $k = 1, 2$,

$$\log(F_{js}/F_{ss}) = \gamma_{js} + \lambda_{js},$$

where $j = 1, \dots, 4$. Under M_3 , $L^2(M_3) = 22.12$; $df = 16$; $P = 0.14$; thus, M_3 is tenable. However, an even better fitting model with the same number of degrees of freedom is

$$\log(F_{jk}/F_{sk}) = \gamma_{jk} + \lambda_j + \delta_k^{c^m},$$

where $j = 1, 2$,

$$\log(F_{jk}/F_{sk}) = \gamma_{jk} + \lambda_j + \delta_k^{c^m},$$

where $j = 3, 4$, $k = 1, 2$, and $\delta_1^{c^m} = \delta_2^{c^m} = 0$,

$$\log(F_{js}/F_{ss}) = \gamma_{js} + \lambda_{js},$$

where $j = 1, \dots, 4$. For this model, hereafter M_4 ($L^2[M_4] = 14.11$; $df = 16$; $P = 0.59$). Model M_4 specifies that origin-destination logit contrasts for nonmanual categories differ between the United States and Britain by an amount $\delta_2^{c^m}$, with estimate $\delta_2^{c^m} = -0.70$ (SE = 0.07). The difference for manual categories, $\delta_1^{c^m}$, is estimated to be slightly larger, $\hat{\delta}_1^{c^m} = -0.86$ (SE = 0.07). Thus, there is some evidence that origin-destination differences in the United States and Britain were similar in pattern but more pronounced in the United States, and they were slightly more pronounced in the manual categories than in the nonmanual categories.

Moving to parsimonious models for the origin and destination distributions of Japan, we consider two models. Both models retain the relevant parts of M_4 for structural mobility in the United States and Britain; that is,

$$\log(F_{jk}/F_{3k}) = \gamma_{jk} + \lambda_j + \delta_k^{\text{am}},$$

where $j = 1, 2$, and $k = 1, 2$,

$$\log(F_{jk}/F_{3k}) = \gamma_{jk} + \lambda_j + \delta_k^{\text{am}},$$

where $j = 3, 4$ and $k = 1, 2$. The first model, M_5 , specifies that all four logit contrasts for Japan are equal:

$$\log(F_{js}/F_{3s}) = \gamma_{js} + \delta_s, \quad (20)$$

where $j = 1, \dots, 4$. The fit of M_5 is relatively poor ($L^2[M_5] = 34.89$; $df = 19$; $P = 0.01$); further, $L^2(M_5) - L^2(M_4) = 34.89 - 14.11 = 23.78$; $df = 3$; $P = 0.00$. However, relaxing the assumption on the fourth logit in (20), that is, letting

$$\log(F_{js}/F_{3s}) = \gamma_{js} + \delta_s,$$

where $j = 1, 2, 3$,

$$\log(F_{4s}/F_{3s}) = \gamma_{4s} + \lambda_{4s},$$

does yield a tenable model (hereafter M_6) ($L^2[M_6] = 15.10$; $df = 18$; $P = 0.66$, and $L^2[M_6] - L^2[M_4] = 15.10 - 14.11 = 0.99$; $df = 2$, $P = 0.61$). Under M_6 , the relative flow from farming into each of the first three categories is the same in Japan, $\hat{\delta}_s = 1.42$ (SE = 0.06), and the flow from farming into the fourth category is slightly greater, $\hat{\delta}_{4s} = 1.80$ (SE = 0.09).

In summary, by using a flexible set of logits that do not necessarily assume a precise ordering to the occupation categories, we have demonstrated that it is possible to develop parsimonious models for characterizing cross-national differences in structural mobility. We have shown that the patterns of structural change were similar for the two Western nations but with a greater degree of change taking place in the United States. Furthermore, the extent to which the change was greater in the United States was parsimoniously summarized in terms of manual versus non-manual categories. While there is significant evidence of structural mobility in Japan, the pattern there is much different from that exhibited by the two Western nations. In Japan, contrary to the patterns observed in the two Western nations, structural mobility was adequately modeled assuming equality across the higher three categories, and the greatest origin-destination contrast was for the logit comparing category 4 to farming.

DISCUSSION

Researchers studying occupational mobility have typically distinguished between structural mobility and circulation (exchange, pure) mobility, but, as demonstrated herein, previous efforts to simultaneously model

these components of mobility are inadequate. This paper couples linear and nonlinear models of association with linear models for logits of marginal distributions, using the new models to simultaneously study the association and the marginal heterogeneity (structural mobility) between origins and destinations. Additionally, we generalize the criteria for combining categories given by Goodman (1981) to account for the case when interest centers not only on association but also on marginal distributions, and in our final example we also characterize the pattern of association within and across societies using models more general than those considered by previous workers.

The methods herein are also applicable in many other substantive contexts. Tables with a one-to-one correspondence between rows and columns frequently arise in longitudinal studies (Hagenaars 1990), in studies of educational and religious homogamy of spouses (Johnson 1980; Luijckx 1994), couple agreement (Hout, Duncan, and Sobel 1987), in studies of educational attainment based on sibling pairs (Mare 1994), and in studies of geographic mobility (Herting, Grusky, and Van Rompaey 1997). Previous workers have modeled these types of tables using log-linear and log-nonlinear models for the frequencies in the table; although the marginal distributions and the discrepancies between these are typically of interest in such contexts, these issues are not readily addressed using such models (e.g., cf. Mare [1994] and Becker and Yang [1998]).

The methods herein are also useful for modeling tables that are not square. For example, Luijckx (1994) considers six tables describing occupational mobility in Hungary over the period 1973–89, where sons are classified into eight occupational categories and their fathers into seven categories. In this case, the methods proposed by Sobel et al. (1985) cannot be used at all. However, our methods can be used to characterize differences in the patterns of association across tables, as well as differences in the univariate marginals across tables.

We hope that the methods used here will prove useful to future workers. At the same time, we recognize that these methods are currently limited in several important respects. In the case of a two-way table, our association models use one set of scores to characterize the row variable and one set of scores to characterize the column variable. In future work, we intend to consider association models with more than one set of row and column scores. Similarly, we intend to consider such generalizations for the case where sets of two-way tables are observed. Second, while our models can handle discrete covariates by treating each covariate class as a table, we intend to take up models where the parameters are allowed to depend on continuous covariates (see Hauser and Grusky [1988] for examples of how to do this in log-linear models). This will enable us to model structural

mobility as a function of economic development, for example. Third, it is useful for some purposes to conceptualize the observed distributions as the results of mixing different probability models (e.g., Mare 1994; Becker and Yang 1998). Combining the results of the present paper with the mixture models (i.e., latent class models) developed in Becker and Yang (1998) also appears to us to be a potentially fruitful undertaking.

APPENDIX

Parameter Estimation

Generalized log-linear models of the type considered herein (e.g., eqq. [8], [9], and [10]) may be expressed as

$$C \log(A\pi) = X\beta, \quad (21)$$

where π is the vector of joint probabilities, A is a matrix of zeroes and ones, C is a matrix of contrasts used to produce the logits in (8) and (9) and the log-odds ratios in (10), X is the model matrix (of full-column rank), and β is the vector of model parameters (e.g., γ_{ab} , λ_{ab} , θ_{ab}). See Becker (1994) for an example of the specification of these matrices and a description of an algorithm due to Lang and Agresti (1994) for fitting such models by the method of maximum likelihood; see also the appendix in Lang and Eliason (1997).

Models that employ a log-multiplicative structure of association (e.g., equation [11]) fall outside of the generalized log-linear model family, but the constrained maximum likelihood estimation algorithm considered in Becker (1994) and Lang and Agresti (1994) can be modified to accommodate the departure from log-linearity. For example, in an $R = C + I$ association model,

$$\theta_{ij} = \theta(u_{i+1} - u_i)(u_{j+1} - u_j) + \text{linear term},$$

and thus some modification is required to accommodate the nonlinear term

$$\theta(u_{i+1} - u_i)(u_{j+1} - u_j).$$

Let $\Phi_l(\omega) = (u_{i+1} - u_i)(u_{j+1} - u_j)$, where $l = j + (I - 1)(i - 1)$ and ω is the vector whose entries are the scores (u_1, \dots, u_I) . The generalized log-linear model is then readily extended to include the log-multiplicative structure of association by adding to $X\beta$ a vector that includes only the subvector $\theta\Phi(\omega)$ for log-odds ratio terms and zeroes for all other entries. Estimation of θ and the scores ω can be accomplished by extending the algorithm for estimation of generalized log-linear models, using a linear approximation to $\Phi(\omega)$ (McCullagh and Nelder 1989, pp. 379–80).

Approximating $\Phi(\omega)$ using a first-order Taylor series expansion about

the current estimate of the scores ω , say $\omega^{(t)}$ for iteration t , and using matrix notation:

$$\Phi(\omega)\theta \approx \left[\Phi(\omega^{(t)}) + \frac{\partial \Phi(\omega)}{\partial \omega'} \bigg|_{\omega^{(t)}} (\omega - \omega^{(t)}) \right] \theta$$

$$= z_1\theta + z_2\eta,$$

where $z_1 = \Phi(\omega^{(t)})$, $z_2 = \partial \Phi(\omega) / \partial \omega' |_{\omega^{(t)}}$, and $\eta = (\omega - \omega^{(t)})\theta$, the iterative algorithm proceeds as follows:

1. Given the current estimates $\omega^{(t)}$ of the scores, use the Lang and Agresti algorithm to estimate the linearized model

$$X\beta + z_1\theta + z_2\eta.$$

2. Check for convergence in the parameters β , θ , and η . If convergence is not obtained, continue to step 3. Otherwise, skip step 3 and continue to step 4.
3. Use $\hat{\theta}^{(t)}$ and $\hat{\eta}^{(t)}$ obtained from step 1 and the relationship

$$\omega^{(t+1)} = \omega^{(t)} + \hat{\eta}^{(t)} \frac{1}{\hat{\theta}^{(t)}}$$

to obtain new estimates of the scores, and then update z_1 and z_2 . Return to step 1.

4. Estimates of the variances and covariances for the parameters β and θ are obtained from the corresponding submatrix of the asymptotic covariance matrix from the Lang and Agresti algorithm. To obtain the submatrix of the covariance matrix corresponding to ω , one further iteration is needed, replacing the nonlinear part of the predictor with:

$$z_1\theta + \hat{\theta}z_2\eta^*,$$

where $\eta^* = (\omega - \omega')$ and ω' is the estimate of ω upon convergence. In this case, $\text{var}(\eta^*) = \text{var}(\omega - \omega') = \text{var}(\omega)$. Also, $\text{cov}(\theta, \eta^*) = \text{cov}[\theta, (\omega - \omega')] = \text{cov}(\theta, \omega)$.

REFERENCES

- Becker, Mark P. 1990a. "Maximum Likelihood Estimation of the RC(M) Association Model." *Applied Statistics* 39:152-67.
- . 1990b. "Quasisymmetric Models for the Analysis of Square Contingency Tables." *Journal of the Royal Statistical Society, Ser. B*, 52:369-78.
- . 1994. "Analysis of Cross-Classifications of Counts Using Models for Marginal

- Distributions: An Application to Trends in Attitudes on Legalized Abortion." Pp. 229-65 in *Sociological Methodology 1994*, edited by P. V. Marsden. Oxford: Basil Blackwell.
- Becker, Mark P., and Cecile C. Balagtas. 1993. "Marginal Modeling of Binary Cross-Over Data." *Biometrics* 49:997-1009.
- Becker, Mark P., and Clifford C. Clogg. 1989. "Analysis of Sets of Two-Way Contingency Tables Using Association Models." *Journal of the American Statistical Association* 84:142-51.
- Becker, Mark P., and Ilsoon Yang. 1998. "Latent Class Marginal Models for Cross-Classifications of Counts." Pp. 293-325 in *Sociological Methodology 1998*, edited by Adrian E. Raftery. Cambridge, Mass.: Blackwell Publishers.
- Boudon, Raymond. 1973. *Mathematical Structures of Social Mobility*. Amsterdam: Elsevier Scientific Publishing Company.
- Clogg, Clifford C. 1981. "Latent Structure Models of Mobility." *American Journal of Sociology* 86:836-68.
- . 1982. "Some Models for the Analysis of Association in Multi-Way Cross-Classifications Having Ordered Categories." *Journal of the American Statistical Association* 77:803-15.
- Duncan, Otis Dudley. 1966. "Methodological Issues in the Analysis of Social Mobility." Pp. 51-97 in *Social Structure and Mobility in Economic Development*, edited by N. Smelser and S. Lipset. Chicago: Aldine.
- . 1979. "How Destination Depends on Origin in the Occupational Mobility Table." *American Journal of Sociology* 84:793-803.
- Featherman, David L., F. Lancaster Jones, and Robert M. Hauser. 1975. "Assumptions of Social Mobility Research in the U.S.: The Case of Occupational Status." *Social Science Research* 4:329-60.
- Glass, D. V., ed. 1954. *Social Mobility in Britain*. London: Routledge & Kegan Paul.
- Goldthorpe, John H. 1980. *Social Mobility and Class Structure in Modern Britain*. Oxford: Clarendon Press.
- Goodman, Leo A. 1979. "Simple Models for the Analysis of Association in Cross-Classifications Having Ordered Categories." *Journal of the American Statistical Association* 74:537-52.
- . 1981. "Criteria for Determining whether Certain Categories in a Cross-Classification Table Should Be Combined, with Special Reference to Occupational Categories in an Occupational Mobility Table." *American Journal of Sociology* 87:612-50.
- Hagenaars, Jacques A. 1990. *Categorical Longitudinal Data: Log-Linear Panel, Trend, and Cohort Analysis*. Newbury Park, Calif.: Sage.
- Hauser, Robert M. 1979. "Some Exploratory Methods for Modeling Mobility Tables and Other Cross-Classified Data." Pp. 413-58 in *Sociological Methodology 1980*, edited by Karl F. Schuessler. San Francisco: Jossey-Bass.
- Hauser, Robert M., P. J. Dickinson, H. P. Travis, and J. N. Koffel. 1975. "Structural Changes in Occupational Mobility among Men in the United States." *American Sociological Review* 40:585-98.
- Hauser, Robert M., and David L. Featherman. 1977. *The Process of Stratification: Trends and Analyses*. New York: Academic Press.
- Hauser, Robert M., and David B. Grusky. 1988. "Cross National Variation in Occupational Distributions, Relative Mobility Changes, and Intergenerational Shifts in Occupational Distributions." *American Sociological Review* 53:723-41.
- Hauser, Robert M., J. N. Koffel, H. P. Travis, and P. J. Dickinson. 1975. "Temporal Change in Occupational Mobility: Evidence for Men in the United States." *American Sociological Review* 40:279-97.
- Herting, Jerald R., David B. Grusky, and Stephen E. Van Rompaey. 1997. "The Social

- Geography of Interstate Mobility and Persistence." *American Sociological Review* 62:267-87.
- Hope, Keith. 1981. "Trends in the Openness of British Society in the Present Century." Pp. 127-70 in *Research in Social Stratification and Mobility 1*, edited by R. Robinson and D. J. Treiman. Greenwich, Conn.: JAI Press.
- . 1982. "Vertical and Non-Vertical Class Mobility in Three Countries." *American Sociological Review* 47:99-113.
- Hout, Michael. 1989. *Following in Father's Footsteps: Social Mobility in Ireland*. Cambridge, Mass.: Harvard University Press.
- Hout, Michael, Otis Dudley Duncan, and Michael E. Sobel. 1987. "Association and Heterogeneity: Structural Models of Similarities and Differences." Pp. 145-84 in *Sociological Methodology 1987*, edited by C. C. Clogg. Washington, D.C.: American Sociological Association.
- Hout, Michael, and Robert M. Hauser. 1992. "Symmetry and Hierarchy in Social Mobility: A Methodological Analysis of the CASMIN Model of Class Mobility." *European Sociological Review* 8:239-66.
- Hutchinson, B. 1958. "Structural and Exchange Mobility in the Assimilation of Immigrants to Brazil." *Population Studies* 12:111-20.
- Johnson, Robert Alan. 1980. *Religious Assortative Marriage in the United States*. New York: Academic Press.
- Kerckhoff, Alan C., Richard T. Campbell, and Idée Winfield-Laired. 1985. "Social Mobility in Great Britain and the United States." *American Journal of Sociology* 91:281-308.
- Lang, Joseph B., and Alan Agresti. 1994. "Simultaneously Modeling Joint and Marginal Distributions of Multivariate Categorical Responses." *Journal of the American Statistical Association* 89:625-32.
- Lang, Joseph B., and Scott R. Eliason. 1997. "The Application of Association-Marginal Models to the Study of Social Mobility." *Sociological Methods and Research* 26:183-212.
- Luijckx, Ruud. 1994. *Comparative Loglinear Analyses of Social Mobility and Heterogeneity*. The Netherlands: Tilburg University Press.
- Mare, Robert D. 1994. "Discrete-Time Bivariate Hazards in Unobserved Heterogeneity: A Partially Observed Contingency Table Approach." Pp. 341-83 in *Sociological Methodology 1994*, edited by P. V. Marsden. Cambridge, Mass.: Blackwell Publishers.
- McCullagh, Peter, and John R. Nelder. 1989. *Generalized Linear Models*, 2d ed. London: Chapman & Hall.
- Miller, S. M. 1960. "Comparative Social Mobility." *Current Sociology* 9:1-89.
- Minick, Susan M. 1997. "Simultaneous Modeling of Marginal Distributions and Log-Multiplicative Associations for Ordered Multivariate Categorical Data." Ph.D. dissertation, Department of Biostatistics, University of Michigan.
- Rogoff, N. 1953. *Recent Trends in Occupational Mobility*. Glencoe, Ill.: Free Press.
- Sinkhorn, R. 1967. "Diagonal Equivalence to Matrices with Prescribed Row and Column Sums." *American Mathematical Monthly* 74:402-5.
- Sobel, Michael E. 1983. "Structural Mobility, Circulation Mobility and the Analysis of Occupational Mobility: A Conceptual Mismatch." *American Sociological Review* 48:721-27.
- . 1988. "Some Models for the Multiway Contingency Table with a One-to-One Correspondence among Categories." Pp. 165-92 in *Sociological Methodology 1988*, edited by C. C. Clogg. Washington, D.C.: American Sociological Association.
- . 1995. "The Analysis of Contingency Tables." Pp. 252-310 in *Handbook of Statistical Modeling for the Social and Behavioral Sciences*, edited by G. Arminger, C. C. Clogg, and M. E. Sobel. New York: Plenum Press.
- Sobel, Michael E., Michael Hout, and Otis Dudley Duncan. 1985. "Exchange, Struc-

- ture, and Symmetry in Occupational Mobility." *American Journal of Sociology* 91: 359-72.
- Ultee, Wout C., and Ruud Luijkx. 1990. "Educational Heterogamy and Father-to-Son Occupational Mobility in 23 Industrial Nations: General Societal Openness or Compensatory Strategies of Reproduction?" *European Sociological Review* 6:125-49.
- Wong, Raymond S. 1990. "Understanding Cross-National Variation in Occupational Mobility." *American Sociological Review* 55:560-73.
- Xie, Yu. 1992. "The Log-Multiplicative Layer Effect Model for Comparing Mobility Tables." *American Sociological Review* 57:380-95.
- Yamaguchi, Kazuo. 1987. "Models for Comparing Mobility Tables: Toward Parsimony and Substance." *American Sociological Review* 52:482-94.
- Yasuda, Saburo. 1964. "A Methodological Inquiry into Social Mobility." *American Sociological Review* 29:16-23.

SYMPOSIUM ON HISTORICAL SOCIOLOGY AND RATIONAL CHOICE THEORY

"We're No Angels": Realism, Rational Choice, and Relationality in Social Science¹

Margaret R. Somers
University of Michigan

An alarm has been sounded that historical sociology is subverting the theoretical aims of social science. Criticizing an array of widely influential scholars, Kiser and Hechter propose that rational choice theory can avoid the trend toward "empiricism" that results from the import of history into sociology. Their position is based on *theoretical realism*—a radically antipositivist thesis that uses ontological and theoretical postulates to theorize about reality beyond positive appearance. A close examination of theoretical realism casts doubts on the epistemological foundations of rational choice theory. *Relational realism*, the alternative introduced here, places greater emphasis on the pragmatic elements of explanation, supporting a more relational, causal-historical, and problem-driven view of theory. A renewed appreciation of what is defined as Kuhn's *historical epistemology* provides the foundation for evaluating these competing research programs.

Almost 40 years ago, in a remarkably prescient first sentence, T. S. Kuhn sent shock waves through our intellectual landscape and forever transformed it: "History, if viewed as a repository for more than anecdote or

¹ This paper originated in a panel on history and theory at the American Sociological Association meetings (1989) where Edgar Kiser, Michael Hechter, and I had a stimulating discussion. Drafts have been improved by the comments of participants at seminars at the University of Chicago and Northwestern University, as well as those from many colleagues along the way, including Julia Adams, Renee Anspach, Mabel Berzin, Fred Block, Paula England, Walter Goldfrank, Martin Hollis, Ira Katznelson, Steven Lukes, John McCormick, Mark Mixruchi, John Padgett, Alessandro Pizzorno, Sonya Rose, William Sewell, Jr., Arpad Szakolczai, Marc Steinberg, George Steinmetz, Arthur Stinchcombe, Wolfgang Streeck, Charles Tilly, Bruce Western, Yu Xie, Mayer Zald, and the *AJS* reviewers. I thank especially Elizabeth Anderson, without whom

chronology, could produce a decisive transformation in the image of science by which we are now possessed" (Kuhn [1962] 1970, p. 1). More recently, in a stinging attack, two influential sociologists issued an alarm about just those social scientists who appear to have embraced Kuhn's challenge most eagerly: "The role of general theory in comparative-historical sociology is under attack" (Kiser and Hechter 1991, p. 1). I juxtapose these contrary perspectives to strike a paradox: Kuhn, a physicist and philosopher-historian of science, turned the prevailing image of science on its head by demonstrating the wrongheaded denial of history in the construction of theory. Thirty years later, Kiser and Hechter *build firmly on the foundations of Kuhn's legacy* to express alarm over the dire consequences of taking up Kuhn's mandate; the profligate use of history, they charge, is subverting theory. This is a paradox that needs explaining: How is it that two historical sociologists can stand on Kuhn's shoulders and denounce the use of history when it was Kuhn himself who changed our understanding of history by demonstrating its centrality in theory and the production of knowledge?

At a time when increasing numbers are taking "the historic turn in the human sciences" (McDonald 1996), this question touches on some of the most critical and contested issues facing not just historical sociology but the social sciences across the board. For one, Kiser and Hechter criticize all research that is not theory driven, or "guided" by "general theory" (1991, pp. 2, 16, 17, 22) and thus make it clear that their specific concerns about historical sociology are less important than their more general calling into question the worth of problem-driven theory building and their skepticism toward alternative research strategies. Second, Kiser and Hechter's view is not an isolated idiosyncratic critique of comparative historical sociology but part of a wide-reaching research program of rational choice theory with many such targets against research too commonly characterized as antitheoretical, relativistic, or empiricist (Green and Shapiro 1994; Friedman 1996). That Kiser and Hechter invoke rational choice not merely to exemplify but as the *only* theory to date that meets their criteria (1991, p. 23) is thus another reason for the general significance of this discussion. Rational choice—and its ancillaries such as game theory and principal-agent theory—is among the fastest growing theories in social science and has itself changed the parameters of recent debate. Exam-

this project would not have been possible. Financial support during the research and writing has been provided generously from the University of Michigan's Rackham Faculty Research Grant and Fellowship, the Office of the Vice President for Research, and the LSA Faculty Assistance Fund, as well as from the European University Institute, Florence, Italy, and the German Marshall Fund. Direct correspondence to Margaret Somers, Department of Sociology, University of Michigan, 500 South State Street, Ann Arbor, Michigan 48109. E-mail: peggs@umich.edu

ining specifically Kiser and Hechter's argument thus serves more generally to clarify the epistemological foundations of one of the most influential theoretical developments of our time. Finally, this discussion transcends a debate among historical sociologists because Kiser and Hechter do not chastise only a few well-known historical sociologists but some of the most highly respected scholars across the social sciences—not just Michael Mann, Theda Skocpol, and Charles Tilly, but also Reinhard Bendix, Kai Erickson, Anthony Giddens, Robert Nisbet, Arthur Stinchcombe, Guy Swanson, Immanuel Wallerstein, and no less than “contemporary survey research in sociology” (1991, p. 16). Surely if this wide spectrum of scholarship is threatening the discipline, we would all do well to take notice.

Kiser and Hechter charge that historical sociology is “empiricist,” wrongly “inductivist,” innocent of the standards for “adequate explanation,” “novelistic,” subversively “relativist,” and “antitheoretical.” Instead of waging a defense, already done forcefully by Quadagno and Knapp (1992) and Skocpol (1994), in this article my primary aim is to turn the tables and examine the foundations of Kiser and Hechter's own position: By what principles of knowledge do they justify their criticisms? And are we convinced by those principles? Since, moreover, Kiser and Hechter's accusations are spun on the terms of historical sociology being antitheoretical, neglectful of mechanisms, too descriptive and interpretative in its use of history, and naive of the necessity for deductivist general theory, then how we evaluate their criticisms fully depends on whether we accept their definitions of history, theory, explanation, and justification. The questions I put to Kiser and Hechter, then, following their own criticisms, follow:

1. What should count as an “adequate” social science explanation? If theorizing causal mechanisms is the criterion for success, and mechanisms are unobservable, how do we know a successful explanation when we cannot see those mechanism?
2. How do we judge among rival theories, and do Kiser and Hechter give us good reasons to believe in *their* theory?
3. And, ultimately, if “general theory” is the criterion for success, as Kiser and Hechter claim, and history the antithesis of theory, what counts as “theory” and “history” in social science?

At stake, I believe, is the epistemological status not just of history but of social science *tout court*. Are theories to be constructed and judged exclusively by theory-driven standards (as per Kiser and Hechter), or (as I insist) should theories be problem driven and judged by grappling with the difficult question of what—beyond the elegance of theory itself—makes an explanation convincing?

In this article I address these issues through a three-part strategy: first, a focus on Kuhn and the Kuhnian legacy; second, a critical focus on Kiser and Hechter (1991); and third, an outline of an alternative to Kiser and Hechter's position. It is in examining Kiser and Hechter's theory-centric justifications for rational choice theory that my argument about the paradox, and the importance, of the Kuhnian influence becomes so important. For rational choice's current epistemological foundations—specifically its insistence on the necessity of, and the justification for, using an exclusively deductivist and exogenous (not subject to problematization) general theory to impute otherwise unobservable causal mechanisms—is part of a trajectory made possible largely by the Kuhnian revolution.¹ I will argue additionally, and again paradoxically, that a slightly different trajectory out of the same Kuhnian corpus provides the most convincing alternative to Kiser and Hechter's model. This alternative is based on a *historical epistemology* (Somers 1996b); it also supports theorizing causality—but with a conception of theory as problem driven, pragmatic, relational, and historical.

It is in considering these contending appropriations of Kuhn in theorizing causality that we confront the paradox that both positions stem from what philosophers call *scientific realism*—the thesis that unobservable phenomena ("theoretical entities") such as atoms and quarks, market forces or individual intentionality, despite the absence of positive (observable) evidence for their existence are "real" enough to be used as explanatory in theoretical accounts. Scientific realism does not fit neatly into either side of the several dichotomies around which Kiser and Hechter frame their attack—between theory and history, between positivism and relativism, between their privileging of deductivism and their denigration of inductivism. But Kiser and Hechter have framed wrongly the terms of debate and their position in it, for these are not the fundamental issues at stake in their argument.

At issue in Kiser and Hechter's rational choice-based attack on historical sociology is a less familiar but no less significant contestation over scientific realism. Induction and deduction are both logics of positivism with confirmation rules based on observation language; rational choice theory, by contrast, rests on a realist thesis, and its method of imputing causal mechanisms from ontological and theoretical postulates makes it

¹ I say the "current" epistemological foundations with full awareness (as I point out *inter alia*) that the roots of rational choice theory go back to 17th-century rationalism, 18th- and 19th-century classical political economy, 20th-century neoclassical economics and decision theory, etc. My argument, however, is not an intellectual history but one addressing the epistemological conditions of possibility for its current resurgence—and these are absolutely post-Kuhnian.

a militantly antipositivist enterprise. All versions of realism accept that causal mechanisms—despite being unobservable—must be used as the basis of explanatory theoretical accounts; but only rational choice realism generates those mechanisms using on “ontic methodology” (Salmon 1984) in which the causal mechanisms of social explanation are postulated a priori from the same general theory that “guides” their research. Thus while scientific realism may seem like an arcane concern to some, any talk of causality (Kiser and Hechter’s as well as my own) that reaches beyond correlation and statistical probability into the territory of mechanisms (“meaningful connection between events as the basic tool of description and analysis” [Coleman 1986, p. 1327]) cannot avoid confronting the challenges of realism.

Questions about what we can legitimately claim to be “real” have, in fact, always been at the epicenter of social science inquiry—not for reasons of metaphysical discourse, but for reasons of explanation and trust in a theory. The crux of it is twofold: First, how can we theorize about—that is, ascribe causality to—that which is unobservable, be it price mechanisms, maximizing preferences, class consciousness, value-driven intentionality, or domination? It is precisely because connective mechanisms are unobservable—unlike correlations of empirical indicators—that positivism has militantly rebuked their inclusion in the realm of scientific theory. That this rebuke has held firm in mainstream conceptions of science is incontrovertible—to wit, the ability of tobacco companies to withstand legal liability on the grounds that no positive evidence has yet been found to prove not just that smoking is associated with but actually *causes* cancer.³

As per Coleman (1990), rational choice rightly rejects this rebuke and rightly insists that conjunction and statistical probability cannot substitute for connective mechanisms in any true account of causality (see also Elster 1989). By insisting on mechanisms, however, rational choice must face up to the second issue at stake in realism, namely the long-standing positivist demand (also rightly justified) for *accountability*: What warrants accepting as true even well-structured and parsimonious explanations that rely on constructs that, since we cannot observe them, we have no reason to accept as true? It was for very good reasons that positivism took on the heavy burden of demanding evidence for theoretical claims;

³ If there is need for any further support for the long-standing epistemological link between the problem of the unobservability characteristic of mechanisms and the difficulty in convincing scientists of true causality, consider that as I write this in October 1996 it is front-page headline news in the *International Herald Tribune* that a “mechanism” has been discovered that actually demonstrates “for the first time” that smoking “causes lung cancer” (emphasis added).

without it, we are susceptible to all and any prejudice passing under the guise of science. Postpositivism's challenge is still justified: Why we should believe in any social science theory—however elegant or parsimonious—that explains by reference to unobservables? This is a question to which we still are owed an answer and to which all theories that traffic in mechanisms must be accountable if they are not to thrive as mere prejudice.⁴

The purpose in bringing attention, then, to what I argue is rational choice's underlying *theoretical realism*—the thesis that belief in an explanation depends on belief in the a priori theory from which it is imputed—is to call its advocates explicitly to account. Kiser and Hechter, in the end, dodge this responsibility: They give us no good reasons why we should believe their version of reality. The advantage of even asking for accountability, however, is that the question may provoke further epistemological deliberation on all sides of the issue; I certainly hope so. In the meantime, my view of rational choice's theoretical realism shares much with Harrison White's rather pugnacious conclusion that rational choice rests on "an underlying ontology of 'spirits' . . . upon angels, that is, upon spirits both disembodied and independent . . . [the] goals or preference orderings . . . essential to . . . rational choice, are appropriate and relevant only to entities which are inertial as well as isolate—angels, in short" (White 1992, p. 301). For those among us who are not angels, however, perhaps we need a realism for the rest of us.

My argument is in part devoted to a reinterpretation of Kuhn's work. It is oddly unfashionable to focus on Kuhn these days—odd since *The Structure of Scientific Revolutions* is "probably the most influential work ever written in the philosophy of science" (Lloyd 1993, p. 207; Fuller 1992, p. 242).⁵ So much has Kuhn passed tacitly into our knowledge culture, however, that his work is sometimes less examined than it is invoked as the crucial point of reference in all discussion of recent transformations in the social sciences—from epistemology to postmodernism. A rethinking of Kuhn and a reconsideration of his legacy, therefore, may be even more important as its explicit retreat from graduate syllabi subjects it increasingly to appropriation without consideration. So I begin with an admittedly unorthodox argument: The legacy of *Structure* is the challenge to rethink theory-construction through the lens of a historical epistemology.

⁴ My thinking about these issues has been enriched enormously through discussions with Martin Hollis.

⁵ Kuhn was not a lone voice among philosophers; he was in the tradition of Duhem ([1906] 1954), Koyre (1957), Quine (1963, 1969), and Polanyi (1958). Why Kuhn broke through to a wider audience is a question for the sociology of knowledge (but see Alexander 1982, Fuller 1992). Kuhn himself famously disclaimed any implications for the social sciences.

The unorthodoxy is purposeful; on the occasion of rational choice theory's appropriation of the Kuhnian legacy, it is fitting to revisit directly the question of Kuhn's significance by exploring his paradoxical duality of influence.

THE KUHNIAN CHALLENGE: BEYOND INDUCTION VERSUS DEDUCTION

When Kuhn published *Structure*, the " 'image of science' by which [philosophers were] . . . possessed" (Kuhn 1970, p. 1) was dominated by positivist epistemology. Epistemology, in the generic, is concerned with how we know, or why certain theories are justified and declared "winners" whereas others barely survive their birth; positivist epistemology, in particular, prescribes rules of logic and observational foundations for distinguishing true knowledge from wrongheaded speculations or mere belief. At the time of Kuhn's intervention, positivists debated over whether induction or deduction was the correct logic for confirming a theory. Kuhn's analysis of science showed the differences between them to be less important than their similarities. Most famously, he pointed to their mutual grounding in the foundational premise of modern science—the sharp delineation between theory and observation: between the self-evident, pre-theoretical existence of empirical reality on the one side, and what are unobserved, hence merely theoretical, "ideas" about reality on the other. For both the logics of induction and deduction, observation is the final adjudicator in confirming or disconfirming a theory.

Drawing from the history of science, Kuhn questioned the theory- and method-independent status of observation. In so doing he brought the term "paradigm" into common academic parlance to signify a worldview that frames not just particular theories but also the questions considered worth asking in the first place and the rules by which they can rationally be answered.⁶ Kuhn showed that what science has considered as confirming evidence has been influenced by what our dominant paradigms allow us to see and, most especially, to care about. This in turn suggested that the debate between induction and deduction could be seen as a false dichotomy held together by a misleading assumption that knowledge progresses by applying the correct logic of theory/evidence relations. Rather than winning or losing on the grounds of logical rules about evidence in confirmation, Kuhn showed those theories that actually have been accepted as explanatory are those that have unseated the available competi-

⁶ "Often the paradigm theory is implicated directly in the design of apparatus able to solve the problem" (Kuhn 1970, p. 27)

tors in the field (1970, pp. 144–45). History shows that it takes a theory to beat a theory.

THE KUHNIAN REVOLUTION: THEORETICAL OR HISTORICAL?

Kuhn proved to be most prescient; the impact of the “Kuhnian revolution” was cataclysmic (e.g., Gutting 1980). The transformation it brought, however, is neither obvious nor singular as Kuhn, like all influential thinkers, has been subject to a multiplicity of appropriations. There is Kuhn 1, the Kuhn accused of “relativism” and “mob psychology”; Kuhn 2, the Kuhn of the history of science and scientific revolutions; Kuhn 3, the theory-centric Kuhn; and Kuhn 4, the Kuhn who outlined a historical epistemology. Kuhn 1 and Kuhn 2, whose ideas are among the most debated topics in the last 25 years of scholarship, are not the focus of this article. Kuhn 3, the chief focus of my critique, is the Kuhn of rational choice theory. Kuhn 4, relatively neglected, is the Kuhn from which rational choice can be challenged and who provides a renewed appreciation for problem-driven research and the centrality of history in the construction of knowledge.

“Theory-centrism” describes the interpretation of Kuhn that focused almost exclusively on his apparent blurring of the fact/theory distinction—a distinction that “positivism made so important . . . it was certain to be denied” (Hacking 1983, p. 170). The instrument of denial was misreading Kuhn to have said that observation is nothing *but* theory, and science was thus exclusively theory driven. But theory-centrism wrongly conflates into one dimension what are three separate aspects of the post-positivist critique. The first is the quite reasonable, but banal, point that no research can take place without some proportion of deductive reasoning, especially to generate new experiments and new measuring instruments. To hold this position, however, is simply to recognize that we go armed with ideas and postulates not solely to test a theory about the phenomena under scrutiny but also to have principles for selecting data or to generate causal analogies (Stinchcombe 1978) that can be tested in turn.⁷ The second extends to the social sciences the problematic argument that since (in natural science) there is often no significant distinction between observable and unobservable entities, they can legitimately be used interchangeably in social theory. The third wrongly extends the justified critique of the possibility of pure theory-independent passive reporting of facts to all research practices not exclusively driven by deductivist logic.

⁷ We also do so to avoid suffering a complete “epistemological crisis” at every decision-juncture in our daily intellectual life (MacIntyre 1980).

The flattening of these into one grand theory-centrism created a gestaltlike change in the image of science: The positivist privileging of "brute fact" was suddenly replaced with a new hegemony of "the primacy of theory" (for correctives, see Pickering 1989, 1992b, 1993; Hacking 1983, 1992; Galison 1987, 1989; Gooding 1992; Gooding, Pinch, and Schaffer 1989). Suddenly it appeared that philosophers had closed the door to reasoning that did not accept what was now an aggressively antiempirical, dogmatic version of deductivist logic—"theory proposes; theory disposes." Overnight, science was rewritten to reflect this new theory-driven story in which experimental practices were demeaned and observations demoted to mere "illustrations."⁸ When such new stories concluded with ordinary positivist morality tales about experiments "confirming the theorists' speculations" they still converged safely with the empirical foundations of deductivism. But it was an easy slide from there to the near-total hegemony of theory in which talk of evidence and reasoning other than theory driven came to be seen as empiricist, naive inductivism. Declared by theory centrists to be nothing more than passive empiricism, experiment—in fact long the driving force in the construction of phenomena and never merely the passive reporting of facts—was now denied an independent role in theory building (Hacking 1983, pp. 159–61, 171). Most ironically, in light of Kuhn's original mandate, history—long the epistemological Other of positivism—became as much a casualty of the new theory-centrism as it had been of pre-Kuhnian philosophy.

RETHINKING THE KUHNIAN LEGACY: A HISTORICAL EPISTEMOLOGY

Kuhn's contribution, I think, should be seen elsewhere—as adumbrated in the first sentence of *Structure*, where he argued against the exclusion

⁸ Hacking, e.g., tells us of radio astronomers Penzias and Wilson who in 1965 experimented on what they thought might be a meaningless phenomenon (because they were not testing a hypothesis) but which held intrinsic interest to them—namely the static found in transatlantic radio. After considerable experimentation, they determined that there was a uniform amount of energy in space. Meanwhile, physicists at Bell Labs were theorizing about what came to be called the Big Bang theory. Only the discovery of the experimental tests—conducted completely independently of any hypotheses about the origins of the earth—could confirm the theorists' speculation of a uniform temperature throughout space. They found it in the work of Penzias and Willson—among the few experimenters in physics to have been given a Nobel Prize. Yet when the story appeared in textbooks it was rewritten to confirm to a theory-dominated one in which the experimenters were represented as merely testing the hypotheses the theorists had already generated (Hacking 1983, pp. 159–61).

of history, not of theory.⁹ He thus posed the challenge of a historical epistemology, a term I use to capture the idea that the history and development of a thing (and not just the logic of its construction) can tell you something fundamental about its nature. The term is purposefully oxymoronic: It intentionally challenges the assumed antihistorical quality of epistemology and instead proposes that all of our knowledge, our logics, our presuppositions, indeed our very reasoning practices, are indelibly (even if obscurely) marked with the signature of time. They are "history laden"—a phrase meant to evoke, to disturb, and to invert the well-known claim that all data are "theory laden"—and to draw attention to the less discussed inverse; namely, that all social and political theory is founded on presuppositional historical claims. In what follows I use the term to cover three neglected but crucially important aspects of the Kuhnian legacy: his implicit focus on *problem-driven* sources of knowledge; his pragmatist notion of *articulation* in research practice; and his demonstration (if not yet theorization) of *path dependency* and *causal narrativity* in explanatory structure.¹⁰

Problem Formation in Theory Construction

As an element of his mandate to go beyond the induction/deduction dichotomy, Kuhn argued that theories are defeated only by competing theories—not just by disconfirming data. But he did not stop there; in keeping with his opening sentence he looked to the history of science to understand why certain theories were even considered candidates for truth in the first place (Kuhn 1970, pp. 144–59). There he found that the most significant

⁹ "The primacy of theory," after all, would hardly in itself be a revolutionary position, always having been the rallying cry of deductivists (e.g., Popper [1934] 1959).

¹⁰ I discuss the concept of a historical epistemology elsewhere at greater length (Somers 1996b; Somers and Gibson 1994). One reviewer makes the suggestion that the term epistemology is too tied up with its own origins in the search for the transhistorical grounds for knowledge to be coupled coherently with *historical*. While I appreciate the accuracy of the comment, a useful distinction nonetheless can be made between epistemology in the generic sense—simply the study or the question of how we determine what counts as knowledge, regardless of how that question is answered—and the specific rules of positivist epistemology that have dominated in standard philosophy (see Rorty 1979). Nothing in my argument suggests any relativism or historicism—in the sense that concepts only have meaning in their original contexts, rather, because I believe that the question of knowledge is a generic question to which all theories must be accountable, it is justified to distinguish a generic use of epistemology from the history of one particular version. I find the conjoining of history and epistemology to be usefully jarring, hence prefacing the term as "purposefully oxymoronic."

factor in determining the spectrum of competing theories was the set of contemporary questions the theories were vying to answer correctly. This laid the groundwork for what was perhaps Kuhn's most important contribution—his challenge to positivism's exclusion of the significance of problem formation in epistemology (Kuhn 1970, pp. 8–9).

Postivist philosophy of science had confined itself to one side of the fundamental distinction between what it called the *logic of justification* and the *context of discovery* (Reichenbach 1951, pp. 229–49; Hempel 1965, pp. 3–18; Popper [1939] 1959, pp. 42–59). Discovery is the context in which we both “discover” those problems deemed significant and generate the initial hypotheses we propose to explain them. It is thus the *generative* moment in which we discover and decide what is worthy of problematizing in the first place. Philosophers considered the context of discovery to be purely subjective (based on psychological propensities of the individual scientist, etc.) and thus of no bearing whatsoever on the validity of a hypothesis. The logic of justification, by contrast, is concerned only with the formal logic of methodology and the correct means for verifying and structuring hypotheses: Are they supported by the evidence? confirmed by experiment? corroborated by stringent testing? (Hacking 1983, p. 6). Since the question of truth and acceptability rested in justification alone, positivist philosophers charged that to confuse discovery with final adjudicating method and logic was a result of “mystical interpretation” (Reichenbach 1951, p. 231) or, even worse, the “genetic fallacy.” That meant wrongly mixing up history—the “origins” of a theory—with its final logical truth.

Kuhn's use of history challenges as indefensible the separation between logic and discovery; doing so, he suggests, actually creates false ideas about how knowledge is constituted and truth decided upon:

I may even seem to have violated the very influential contemporary distinction between “the context of discovery” and the “context of justification.” Can anything more than profound confusion be indicated by [this] admixture of diverse fields and concerns? Having been weaned intellectually on these distinctions and others like them, I could scarcely be more aware of their import and force. For many years I took them to be about the nature of knowledge. . . . Yet my attempts to apply, even *grosso modo*, to the actual situations in which knowledge is gained, accepted, and assimilated have made them seem extraordinarily problematic. (Kuhn 1970, pp. 8–9)

Once history shows us that theories are tested against competing theories, Kuhn argues that the logic/discovery distinction proves untenable. Instead, it is the historical context of discovery that determines the spectrum of theories available for comparison at any point in time. Why, for exam-

ple, at T_2 rather than at T_1 , does theory A win out over the field of competitors {A,B,C . . . }? Kuhn shows that for historical reasons theory A could appear to be justified in the field of theories {A,B,C . . . } available for comparison, but look not at all justified if the field of available theories is or was {A,D,F} (1970, pp. 23, 94–110, 144–59).

Two key insights can be extrapolated from Kuhn's focus on problem formation. First, because scientists ask different questions at different points in time—often because of the questions that have already been asked—different answers also satisfy them at different points in time. Theory A may beat B at time T_1 but lose to B at time T_2 because those questions considered historically compelling may have changed (1970, pp. 102–6). Thus, for example, Galileo's science was accepted at the exact moment when changes in artillery made it urgent to figure out where cannonballs would land and how far they would travel with a given amount of gunpowder. Historical exigencies most commonly bring certain questions into existence in the first place. "Civilizational concerns" (Zald 1991) post-1989, for example, explain why social scientists now problematize democratization at a level of intensity completely absent only 10 years earlier, thus bringing new candidates (e.g., Putnam 1993) or recycling ones once thought moribund (e.g., social contractarianism) into the field of contending theories. This demonstrates that contained within all theory is the *temporality of the question*, and a confirmed theory at T_1 may well be subject to future disconfirmation at T_2 when pressed by a different kind of problem (Kuhn 1970, pp. 198–200, 209–10). This does not suggest that there is no best theory—of course, there is—rather, that a certain humility is necessary in the face of what is not an absolute triumph but an "epistemological gain" (Taylor 1989; Longino 1990).

Kuhn anticipated that skeptics would reply that it is an accident when we find the best theory because, if the full field of logically possible theories {A,B,C,D,F} had been up for comparison, there could have been only one winner; furthermore, that one winner would have won in relation to any subset of those theories (1970, pp. 171–72). An answer to these criticisms can be drawn from the second major implication of Kuhn's focus on problem formation and the historicity of scientific justification: The adjudicative status of facts and evidence depends upon the temporal context in which it is discovered; indeed the epistemological significance (degree of importance in confirming or rejecting a theory) of the same pieces of fact and evidence could be completely at odds depending upon the historical moment and context. For example, potentially disconfirming facts or "anomalies"—pieces of evidence that are incorrigible vis-à-vis a theory/prediction—vary in their importance depending on the point in time the data are observed. In a period of "normal science" an anomaly is treated

not as a refutation of the theory but as an invitation to further develop the theory to accommodate the observation or, alternatively, to sweep it under the rug, sight unseen. In a moment of scientific "crisis," however, the very same piece of evidence can assume an entirely different historical significance and can be used to refute the theory, frequently creating confusion and disarray among its former defenders (Kuhn 1970, pp. 97–98). The status of anomalies is also affected by the historical discovery of new anomalies: If a theory is suddenly inundated by an onslaught of new anomalies some previous anomaly—once considered minor or merely a "problem of measurement"—may suddenly be used as major evidence to show how terribly a theory has failed (pp. 52–91).

Contrary to the prescriptive rules that exclude history from the criteria for the acceptance of knowledge, Kuhn's historical investigations suggest that theories are rarely adjudicated primarily on the basis of the logics of justification—whether inductive or deductive (1970, pp. 3, 26, 94–95). Rather, the historical processes of problem formation appear to be driving knowledge: Only in response to a particular question are facts transformed into evidence at all, and that a particular question was being asked was itself a matter of historical moment (Kuhn 1970, p. 103; Collingwood [1939] 1970; Somers 1996b). Standard epistemology declared that matters of truth are decided by purely logical procedures that are unaffected by historical circumstance; subject-matter data may be historical, but the foundations for knowledge are themselves outside of time. Yet here was Kuhn showing that in practice the criteria for justification have varied historically: Why a theory is confirmed at one time and not another requires a historical and causal explanation, rather than a strictly logical one, as "changes in the standards governing permissible problems, concepts, and explanations can transform a science" (Kuhn 1970, p. 106).

Articulation in Research Practice

Kuhn's demonstration of the epistemological centrality of problem formation and his critique of the induction/deduction dichotomy show the great failure of the theory-centric reading of Kuhn to distinguish between "empiricism" and the altogether different practice of experimental activities. In its conflation of experiment into a passive empiricism, theory-centrism mistakenly combined two separate activities—"reporting" (via passive representation of observations) and "doing" and "causing" (via intervening; see Hacking 1983, p. 173; Humphreys 1988, 1989)—into the single and justifiably vulnerable concept of passive accumulation of data. But Kuhn was equally critical of claims for disproportionately deductivist

logic, for history also demonstrates that the relative import of general theory-driven testing has been radically exaggerated (1970, pp. 23–33). The reason for this is simple: “There are seldom many areas in which a scientific theory *can* be directly compared with evidence” (p. 26; emphasis added). Science reveals many avenues to pathbreaking knowledge; testing theories by observing only the phenomena designated by a preexisting hypothesis is only one such avenue.

Indeed Kuhn suggested that the most common type of research practice was what he called *articulation*: “A paradigm [theory] is rarely an object for replication. Instead, like an accepted judicial decision in the common law, it is an object for articulation and specification under new or more stringent conditions” (1970, p. 23). The analogy to the common law is telling, for like Oliver Wendell Holmes’s ([1881] 1963, p. 1) famous excoriation that law is made by judges rather than applied through logic (the “life of the law is not logic but experience”), Kuhn is suggesting that new theory is constructed not through applying the logic of deduction but through articulation—the pragmatic interaction of theory with historically changing problems. In this effort to capture the practice centeredness of the research process through the term articulation, Kuhn, like Holmes, is challenging the priority of abstract rules of logic over the historical construction of knowledge—apparently building from the original Greek meaning of *historia* as practical “inquiry” (*Oxford English Dictionary*, 4th ed., s.v. “history”). Experimenting through articulation “can resemble exploration” (Kuhn 1970, p. 29), “more than any other sort of normal research, the problems of paradigm articulation are simultaneously theoretical and experimental” (p. 33). Articulation includes “resolving some of its [the paradigm-theory’s] residual ambiguities and permitting the solution of problems to which it had previously only drawn attention . . . manipulations of theory undertaken, not because the predictions in which they result are intrinsically valuable, but because they can be confronted directly with experiment” (pp. 27, 30). In a direct challenge to the possible hegemony of either induction or deduction, he stresses that “the need for work of this sort [articulation] arises from the *immense difficulties* often encountered in developing points of contact between a theory and nature” (p. 30; emphasis added). Articulation can also radically transform accepted logics. In Galileo’s time, for example, previously unencountered historical and practical exigencies induced methodological instabilities; where scientists were once satisfied with qualitative explanations, they may suddenly have demanded accurate quantitative predictions with the introduction of cannonballs. Kuhn’s discussion of the pragmatics of articulation was brief and only suggestive; the challenge he posed was to take this suggestive notion and nurture it, as per Hacking (1983) who returns to an “old-

fashioned concept of history, as history not of what we think but of what we do" (p. 17; see also Pickering 1992a).

Kuhn's Explanatory Challenge

Kuhn is usually read as challenging only the logical rules of theory confirmation—hence the wrongheaded association of Kuhn with “irrationality” and “relativism” in science (e.g., Lakatos 1970). But recall that Kuhn demonstrated that the criteria for theory acceptance varied throughout history. He did so by constructing an empirical causal explanation that gave a convincing alternative theory for how and why science developed the way that it did. Thus while he did not explicitly prescribe an alternative set of rules for the correct structure of an adequate explanation, Kuhn’s explanatory account of Western science is so powerful that I believe it can be read as a *performative* demonstration and an answer to epistemology’s central methodological question: How does a set of hypotheses, if true, explain why something happened?

At the time of Kuhn’s writing philosophers almost uniformly advocated the deductive-nominological (D-N) method, or the “covering law” model of explanation, in which an explanation is derived as an instance of a conditional proposition stating empirical regularities (Hempel [1942] 1959, 1962, pp. 247–48; Popper 1959; Reichenbach 1951; Miller 1987; Outhwaite 1987; Salmon 1984).¹¹ How something happens must follow deductively from the conjunction of sentences describing laws and initial conditions, which is expressed in the logical model of those conditions. Because theories are “covered” by deductive laws (hence the “covering law” model), prediction and explanation are considered to be subsumed under the same logical process; explaining the past and predicting the future involves the same operation. Most notable, and an easy source of confusion, about the covering law model is its emphatic rejection of causation unless subsumed under nomothetic (general and universal) law. Causation in this sense refers *not* to laws based on event regularities but to explanation of the specific mechanisms that cause X to lead to Y. According to a strict positive logic derived from Hume (1938), theorizing a true causal explanation based on spatiotemporal connective mechanisms is nonscientific because actual causal pathways can never be positively observed; only correlations or regular conjunctions of events can be empirically grounded in observation and thus supported by scientific law. By rejecting the scientific status of underlying spatiotemporal causal mechanisms (because they are neither lawlike nor empirical), classical positivism was able to combine under a

¹¹ Inductive-statistical methodologies share a commitment to the deductive form of the covering law model of explanation

single rubric of general law both prediction and explanation. But explanation in this epistemology is not an explanatory account of *how* something was actually caused; instead, it is a limited deduction from empirical regularities (Hempel 1959, 1965, p. 359).¹²

Because it subsumes explanation under the rubric of predictive universal laws, the covering law model cannot disclose the underlying causal mechanisms underwriting chains of events (e.g., "enchainment"; Abbott 1993, 1995; Coleman 1986, p. 1327); it cannot allow for the contingency of outcomes or explain how temporal sequences, conjunctures, and spatial patterns matter to theory construction (e.g., Aminzade 1992; Griffin 1992, 1993); it searches only for similarities, not variations and processes (Tilly 1995c). The result is that time and space become an empty "picture frame" (the scope parameters) in which the real action—now assigned to laws—takes place. But the results can be nonsensical: "Can Newtonian dynamics really be *derived* from relativistic dynamics?" (Kuhn 1970, p. 101; original emphasis).

Kuhn's explanatory account of Western science can, by contrast, serve as an implicit template for understanding how causal explanations, to be effective, must be temporally and narratively constructed. By giving us a deep causal account—full of causal mechanisms—of the development of science, he introduced temporality and historicity as category of explanation and demonstrated the limits to the covering law model. A successful explanation, as Kuhn shows us, one intended to account for *how* and *why* scientific knowledge developed, could not be constructed through logical deductive entailment under universal laws but only through a causal historical narrative establishing causal chains and relational structures. A given theory not only has a history, but is a history—at each stage it bears the sedimentation and residue of its previous history, of a series of encounters with confirming or anomalous evidence (MacIntyre 1980). Kuhn described what he considered to be the teleological logic of the covering law approach as "stages *toward* a set goal, a permanent scientific truth" (1970, p. 173). His own causal historical account of science, by contrast, he depicted analogously to evolution as contingent, path dependent, and driven by spatiotemporal causal mechanisms and linkages—not by the "goal" of a universal law: "We are all deeply accustomed to

¹² Although its strict epistemological legitimacy has been considerably diminished in recent philosophical argument (Salmon 1988; Miller 1987; Cartwright 1983; Outhwaite 1988), as Steinmetz (1998) points out, a "watered-down" version is still dominant within American sociology. Its methods courses, statistics textbooks, and articles in leading journals all point to the same basic concept of theory as "constant conjunctions of events" and empirical generalizations expressed as universal statements of the covering law type (Steinmetz, 1998, p. 172); see esp. King, Keohane, and Verba (1992), Turner (1992), Wacquant (1993).

seeing science as the one enterprise that draws constantly nearer to some goal set by nature in advance. But . . . *Origin of Species* recognized no goal set either by God or nature. . . . Even the eye and hand of man . . . were products of a process that moved steadily *from* primitive beginnings but *toward* no goal . . . the evolution of scientific ideas . . . may have occurred, as we now suppose biological evolution did, without benefit of a set goal, a permanent fixed scientific truth, of which each stage in the development of scientific knowledge is a better exemplar" (pp. 171–73; emphasis in original).

Kuhn's analysis gave new life to an old suspicion: That an explanation intended to account for how and why something happened—and not one generated only for predictive purposes—must be established through causal chains and mechanisms.¹¹ Absent these linkages, the covering law model undermines its capacity to present a true explanation (Cartwright 1983, 1989). Striking implications for the epistemological status of temporality and process can thus be derived from Kuhn's demonstration that a successful explanation cannot work through logical deductive entailment. Most notably, it makes clear that there are alternatives to the covering law model that cannot be dismissed as nontheoretical, particularizing, and "storytelling"—that is, "merely historical." Put slightly differently, alternative methods such as causal narrativity and path dependence *are* historical, and Kuhn has given history a new and long overdue significance and epistemological stature in the construction of theory.

WHAT IS HISTORY?

When Kuhn wrote *Structure*, the dominant image of history in science as merely "anecdote or chronology" (Kuhn 1970, p. 1) was mirrored in the social sciences, long divided between interpretism and positivism (Apel 1984; Hollinger 1980). Although apparently at complete odds, arguably what kept them joined in battle for so long were mutually caricatured depictions of each other. For both positivists and interpretivists, science was a matter of general laws and logical principles. For positivists that was its appeal; for interpretivists, the grounds for rejecting its applicability to the human sciences. Similarly, for positivists and interpretivists alike, history was historicist—hence beyond causal analysis. A historical epistemology implicitly knocks out the foundations of these very shared assumptions: Neither science nor history fit the images depicted by either side of the debate. Science does not exist outside its own historical condi-

¹¹ Gould (1989) and MacIntyre (1980) each capture the elegance of this approach, albeit in very different ways.

tions; those historical conditions, however, are themselves constituted by theoretical causal sequences and relational structures.

A historical epistemology thus invites us to think about history as a dimension of epistemology and thus as a constituent, rather than an illustrative, element in the conditions and the construction of knowledge: "History, we too often say, is a purely descriptive discipline. . . . [But] how could history of science fail to be a source of phenomena to which theories about knowledge may legitimately be asked to apply?" (Kuhn 1970, pp. 8, 9). Like Stinchcombe (1978, 1984) earlier, Tilly has recently taken up this invitation in his use of the thick/thin metaphor to characterize as "thin" the use of history as no more than a "transparent medium carrying along more substantial causes," whereas in "thick" history time is "drenched with causes that inhere in sequence, accumulation, contingency, and proximity" (1994, p. 270). Thick history is a causal participant in the construction of knowledge without which theory cannot explain the world. The implications of this expanded sense of history from thin to thick, from passive medium into the realm of epistemology, show not only that theories are inherently historical, but also that what we call "history" is inherently theoretical. Indeed it is the "thin" use of history that has "repeatedly led social scientists to the mistaken conclusion that historians are particularizers while social scientists are generalizers" (p. 270). Tilly thus pays tribute to Kuhn who first defrocked these caricatures and challenged us to develop epistemologies that no longer thrive on either one: Would we continue the work of liberating history from its "thin" status as "antitheoretical," "purely descriptive," and "particularizing"? Or would we fall back on easy prejudices and "an entire arsenal of [false] dichotomies" (Kuhn 1970, p. 8)?

The Failure to Meet Kuhn's Challenge

Kiser and Hechter have taken the latter route. In their attack on historical sociology they rely relentlessly on the very caricatures Kuhn so forcefully challenged us to move beyond. Historical sociology is accused of (1) using naive inductivism rather than deductivism, and (2) doing "merely" descriptive narrative history rather than "general theory." At first glance, it appears as if Kiser and Hechter are simply rehearsing an orthodox positivism, criticizing historical sociology for violating rules of logic and not producing adequately lawlike generalizations. That they are, however, forcefully *antipositivist* in their commitment to rational choice's *theory-driven* logic of social analysis and causal mechanism, makes it clear that their position is more complex and more challenging. To assess the justification for this epistemological mix, we need to place Kiser and Hechter's position in the context of scientific realism.

TWO REALISMS OUT OF KUHN: THEORETICAL AND RELATIONAL

One of the great paradoxes of the Kuhnian revolution is that it gave new life to an old idea. Scientific realism is the centuries' old rationalist thesis that objects of knowledge exist and act independently of our direct knowledge of them, even when they are unobservable "theoretical entities," such as electrons and atoms, or social structures, classes, and market forces (Bhaskar 1979, 1986, 1989; Hacking 1983; Putnam 1975). To understand why so much hangs on this claim, it is worth remembering that the issue of observability has been a thorn in the heart of science's basic aspirations since the 17th century. On the one hand, since Hume, science has insisted that for any knowledge claims to qualify as sound there must be sensory empirical proof; hence Comte's, Mill's, and Durkheim's insistence on a positive sociology. At the same time, however, since Descartes science has also had the rationalist goal of understanding laws, forces, and structures that, even though beyond the senses, are nonetheless believed to be the *real* forces at work in the world—albeit a reality beyond appearance.¹⁴ The tension, of course, is that prioritizing positive evidence, versus rationalist or *realist* postulates, are competing goals. Do Durkheim's "social facts" really exist—that is, explain—anything? We can observe the (apparent) results of racial discrimination, but how do we prove a causal claim about racial characteristics? Are Marx's economic laws really true? We can count income distribution, for example, but how do we prove a causal claim about class exploitation? Are people really rational choosers? We can count low voter turnout, but how do we prove that the cause is really utility maximization or free ridership, as opposed, say, to registration difficulties put in the way of voting by state and county governments?

Although rationalism and positivism long vied for dominance, in the 20th century the victory had gone decisively to the logical positivists who insisted that knowledge had to be justified by what was called "theory-independent observation language"—that is, theory must hold up against what in the last instance can be observed empirically, not what might (in theory) really be deeper, unobservable forces. Counterintuitively, this positivist position is *antirealist* because it does not define as "real" anything beyond the observable (correlated with indicators or accessible to inference and measurement). This is not to say that antirealists accept no such thing as reality; simply that what is inaccessible is—for all *scientific* purposes—as good as "not-real," since we have no empirical support for believing in it. Realists call this "actualism" as it "denies the existence of underlying structures which determine . . . events, and instead locates the

¹⁴ "Newton saw apples fall with his eyes but the force and law of gravity are not to be perceived" (Hollis 1994, p. 4).

succession of cause and effect at the level of events [empirical regularities]" (Collier 1994, p. 7).

At stake in the question of realism is thus a great deal more than talk of metaphysics; these are questions of what counts as knowledge, how to build an explanation, and what it takes to justify an explanation based upon unseen causes. Questions about the reality of unobservables clearly matter because most social theories attribute causality to—that is, theorize about—unobservable phenomena, for example, explaining the fall of the Berlin Wall by reference to a democratic “political culture.” If theories can only be justified empirically, then what counts as a theory of democratization must be limited to inferences from statistical regularities about, say, opinion-pollled discontent and degrees of capital infusion. To be sure, attitudinal indicators and factor analysis can be used to translate unobservable theoretical entities into correlations; but correlations are not causes. So while we can count survey answers and numbers of firms, a force called “political culture” still remains inaccessible to the senses and so—for antirealists—cannot be attributed the status of a true cause.

Antirealism thus puts strict limits to what we can justifiably claim to be able to explain, grounding its criteria on what we can know positively to be real. Realism, by contrast, rejects these limits to what counts as theory and is the inverse in every respect—beginning with its postulate that little about the social or natural world makes sense *unless* one believes in the reality of unobservable forces. From this premise follows the difference over what counts as theory: Whereas antirealists limit explanation to the covering law model and exclude causality, realists insist that explanation can and must be causal in the sense of accounting for mechanisms, regardless of empirical access. Realism is clearly attractive because it allows us to extend the reach of theory beyond the limits set by positivism and to theorize about that which positivism excludes—the causal power of unobservable mechanisms.

By now it should be obvious just how important the Kuhnian revolution would have been for a rebirth of realism. Kuhn’s historical challenge to the absolute primacy of observation in adjudicating scientific truth gave new life to the premise that there is more to reality than what is observable or easily translated into empirical indicators. In fact, Kuhn’s demonstration that scientific theories are rarely disposed of due to lack of fit with data, and more often because of the appearance on the scene of a more powerful (more problem-generating) competing theory-paradigm, suggested that it is not only possible but justified to claim that theories about reality (rather than observation) were actually the prime movers of scientific knowledge. How else was it possible or justified for Newton to use the theory of gravity to explain a falling apple? At the end of the day, after all, despite all the strains and struggles to make the consequences

or effects of an unobservable observable (preferably "calibratable"), gravity itself was and is unseen and thus remains a theoretical entity.

The monumental challenge that post-Kuhnian realism continues to face, however, is how to convince us why we should believe in an unobservable reality's putative existence. It is one thing to agree instinctively that there exist social and natural forces and dynamics beyond the access of empirical perception and instrumentation; it is wholly another thing to consider as more than prejudice any one theory about the nature of those social realities. Thus, for example, if faced with a theory that explains the collapse of the Soviet Union in terms of market forces, game-theoretic agency, or political culture, the antirealist asks, Why should anyone believe these claims to be true when they attach causal power to phenomena that we cannot see, feel, hear, or touch? As long as we continue to believe in theories dealing in unobservable phenomena (as in most social science), it is fair to ask for an account of how we can believe—reliably—in such causes (Hollis 1994, p. 12).

It is on this issue of epistemic access that realism is radically divided; and I want to suggest this division can be interpreted as two contrasting routes out of Kuhn, each bearing differing traces of a contested Kuhnian legacy that has become part of our contemporary knowledge culture. On the one side runs *theoretical realism*. Following the theory-centric route, theoretical realism assumes that theories survive to the extent that they are true pictures of what is real, that this truth is guided by a general theory, channeled through deductive logic, and in the final instance confirmed by the same general theory—in this case defined as a set of ontological axioms about the unobservable generative mechanisms at work in the social world. Harking back to 17th-century Cartesian rationalism, building in part on Friedman's (1953) deductivist method of "positive economics,"¹⁵ and bolstered by a theory-centric reading of Kuhn, theoretical realism posits theories to be notoriously "underdetermined by data" (Newton-Smith [1958] 1978, p. 72; Boyd 1973). Lakatos (1970, 1976, 1978), for example, perhaps the most influential representative of theoretical realism (and certainly the primary influence on Kiser and Hechter), represents himself as having improved on Popper in light of (a theory-centric) Kuhn: Competing theories should be adjudicated not on the basis of which one displays consistency with evidence, but on the basis of which one *outpaces* the evidence: The—well planned—building of pigeon holes must proceed must faster than the recording of facts which are to be housed in them (Lakatos 1970, p. 100).

¹⁵ I discuss below, however, the ways in which rational choice's theoretical realism with its belief in the truth of theories differs significantly from Friedman's instrumental "as if" approach to theoretical assumptions

Lakatos describes the scientific project not as a paradigm but as a "research program" comprised of an exogenous "hard core" not subject to debate, problematization, or disconfirmation that, in practice, takes on the status of a metaphysic; this is what Kiser and Hechter and many rational choice theorists consider a general theory. It is protected by an outer "auxiliary protective belt" which receives implications/predictions from the hard core, translates them into hypotheses, and allows *only* these to be tested or to be open to new formulations and measurements (Lakatos 1978). The inner hard core, by contrast, remains pristine and fully protected from threats of "naive falsification." Thus, as we will see, rational choice theorists protect their general theoretical assumptions about individual utility maximization and intentionality-driven mechanisms within the hard core allowing only the implications for a specific case to be sent out into the potentially disconfirming world of testing.

Against "empiricist epistemology," theoretical realism thus follows Lakatos and assigns priority to a philosophically based framework of research from which explanation is logically inferred. In marked contrast to Friedman's explicitly instrumental use of "as if" assumptions to generate testable hypotheses, theoretical realism generates hypotheses from "a prior argument about social reality [which] should operate . . . as a framework for research and as a *regulative principle akin to the principle of truth*" (Lloyd 1986, p. 9; emphasis added). And in contrast to the standard positivist criteria for causality (correlations of indicators in the form of manifest events) theoretical realism rejects "the necessity of regular mechanical connection between events in favor of the idea of the *essential* causal powers of kinds of things" (p. 8; emphasis added). For theoretical realism the work of causal analysis is to "discover," at the most "general" level, the structures, powers, propensities, and liabilities of both persons and social structures so that general, lawlike statements about both of them can be made, statements that refer to the "essential but unobservable powers and tendencies of natural kinds" rather than to event-event regularities (p. 156). This conception of general theory thus marks theoretical realism's departure from both positivism and the historical post-Kuhnians: theory consists of identifying the essential and timeless properties of the social and natural world that will constitute its Lakatosian hard core.

Over the last decade, what can be fairly dubbed a *relational* and *pragmatic* realism has emerged as a post-Kuhnian philosophy of science that is squarely on the other side of the epistemological divide from theoretical realism.¹⁶ Relational realism posits that belief in the causal power of unob-

¹⁶ Steinmetz (1998) has made the criticism that he sees no compelling reason why Bhaskar's "critical realism" is not an adequate term for the kind of realism I call "relational and pragmatic." Although I generally do not endorse the proliferation of

servables—such as states, markets, or social classes—does not depend on the rationality or truth of any given theory but upon practical evidence of its causal impact on the relationships in which it is embedded. Following the historical and more pragmatist reading of Kuhn, relational realists believe that, while it is justifiable to theorize about unobservables, any particular theory entailing theoretical phenomena is historically provisional. For relational realism that means one can believe in the reality of a phenomenon without necessarily believing in the absolute truth or ultimate reality of any single theory that claims to explain it. Belief in a phenomenon or an outcome instead depends on evidence of its causal, practical, and relational significance in time and in space, a practice that often entails building models to represent theoretical entities and testing them by observing effects and inferring by “abduction” to the existence of the entities. But the practices need not be so formal: Why do relational realists continue to believe in electrons when theory X has been proved wrong? Because, regardless of the successive fate of competing theories of electrons, and long after the current theory has been surpassed, supermarket doors will still open automatically when we step on the rubber mat—thanks to whatever reality the theoretical concepts of electron, and photon now represent (Hacking 1990b, p. 356). Substitute the term society, domination, gender, or any variety of social concept, and the same argument holds. For example, a relational realist would use pragmatist reasoning to argue that despite the fate or fashion of any particular theoretical concept such as “sex roles,” “sexual division of labor,” or “gender,”—each of which represents a different causal conception of an unobservable postulated reality—we have reason to believe in the causal force of that which these terms variously attempt to signify largely for one reason: When we dress a baby in blue, we can observe that people treat that baby differently than when we dress that same baby in pink.

Relational realism is thus a “minimalist” realism (Humphreys 1988; Longino 1990) in that it presumes that if one is going to be a realist at all—that is, assign mind-independent status to elements of the world—then, by definition (and humility), one must be agnostic about the absolute truth of any given theory about the world. This is what Bhaskar means when he distinguishes between the “intransitive objects of science . . . and the changing (and theoretically-imbued) transitive objects which are produced within science” (1986, p. 52). Belief in an entity or a phenomenon may well outlast numerous conflicting and failed versions of a theory—to wit, the perdurability of the belief in “society” despite a surfeit of failed

neologisms, in this case I find myself enough at odds with Bhaskar’s views (that causality is found in an entity’s *essential properties*) to require an alternate realism built on relational and pragmatist premises

versions of social theory. Where the two realisms differ, then, is that while theoretical realism attributes an ontological truth to the *theoretical* phenomenon (e.g., the theory of electrons or the theory of market equilibrium), relational realism focuses on the relational *effect* of the phenomenon itself (e.g., the impact of the hypothesized electron on its environment or of the hypothesized market forces on an observable outcome).

Relational realism is pragmatic and relational because it believes phenomena to have causal properties only in virtue of relational evidence, and because such knowledge can only emerge through "the temporality of practice" (Pickering 1992a, pp. 2–3, 9; see also Camic and Xie 1994). It is the relational effect of a hypothesized phenomenon such as gender, sex roles, on a problematized outcome that makes us believe in its causal power and not because of any single theory. To be sure, regardless of whether they are social or natural scientists, relational realists believe in the importance of determining which theories more closely represent reality. But the relational realist does not base her preference for one causal claim over another on whether or not she has found a theory more lean and parsimonious than another. Science has shown us that some theoretical reality claims are false just because the entities in question are not really there—whatever the logical validity, beauty, or parsimony of the theories that originally invented them: Epicurean atomism is not true, there are no humors, nothing with negative weight exists, women's wombs do not wander, "phlogiston is one with the witches and the dragons" (MacIntyre 1980, p. 72). But other causal claims have endured despite a succession of different theoretical positions to account for them. Social phenomena endure; but the "theoretical entities" that have purported to explain them are socially constructed—some more convincingly than others because they are more pragmatic and relational.

What's Wrong with Theoretical Realism?

Kiser and Hechter are theoretical realists—and thus heirs of the unbalanced, theory-centric reading of Kuhn. My criticism, however, centers not just on their theoretical realism, but on their peculiar mixture of pre-Kuhnian antirealism in their dedication to the covering law model of universal regularities, with their postpositivist theory-centrism as expressed in their commitment to causal mechanisms. Each in itself is problematic, but in combination they operate in such tension that they generate incoherent grounds for critique.

"WHAT ADEQUATE EXPLANATIONS MUST ENTAIL"

Kiser and Hechter start out on the wrong foot: "Wide agreement exists across social science" about the "first requirement of an adequate explana-

tion—causality” (1991, p. 4). Contra Kiser and Hechter, no such consensus exists. Contestation over causality has been the greatest thorn in epistemology ever since Hume convinced philosophers that the most we can call an explanation is a statement of correlations, and the most we can say about correlations is that they are driven not by necessity (as with mechanisms) but by statistical probability—which does not claim to be a true account of the causal mechanisms that actually produce the observed effect. As any frustrated social scientist knows from the difficulties of trying to prove causality despite overwhelmingly strong statistical correlation between, say, capital punishment and black defendants/white victims, the battle to overcome the stigma of what classical positivist skepticism from Comte, to Mill, to Popper has considered the mere “psychological appeal” of “ideas” about causation (beyond conjunction) has by no mean been won (Popper 1950, p. 722).¹⁷

Underlying these questions of method, as I stressed above, is positivism’s antirealist epistemology: What counts as an explanation is limited by what we can claim to know empirically to be real enough to refer to as a cause in the first place. Questions of explanatory methodology are thus inexorably linked to deeply contested issues over realism and to difficult questions about how to explain by reference to causal mechanisms whose reality is based on forces that can neither be observed, nor manifested beyond correlated indicators, nor conjured up *a priori*. What “adequate explanation must entail” is thus hardly a matter of consensus.

Causal Relations and Causal Mechanisms

Seemingly unaware of these ambushes, Kiser and Hechter plunge headlong into this methodological thicket, asserting flatly—and wrongly—the existence of similar agreement over the structure of an adequate explanation, namely both causal relations and mechanisms (1991, p. 4). “Causal relations” refers here to the standard covering law model in which a true hypothesis must be analyzed logically in relationship to that which is to be explained as a generalization of constant conjunction: “In essence, causal explanation works by subsuming events under causal laws (Elster 1983, p. 26)” (Kiser and Hechter 1991, p. 6). With no attention to the fact that the covering law model flatly rejects explaining by causal mechanisms, they demand just that: “A complete explanation also must specify a *mechanism* that describes the process by which one variable influences the other” (p. 5). Hence the first glaring incoherence: Kiser and Hechter at-

¹⁷ Bertrand Russell once described the law of causality as “a relic of a bygone age” (quoted in Entrikin 1991, p. 110); Popper dismissed Weber as naive for not recognizing that causality can only be a feature of universal laws (1959, p. 722).

tempt to combine a covering law model logic based on empirical regularities with a method of imputed causal mechanisms—exactly that which covering law model theorists reject as part of a legitimate explanation.¹⁸

Here Kiser and Hechter stumble badly. Deep into the terrain of what their original authority (Hume) declared metaphysics and theology (realist views about causal mechanisms), Kiser and Hechter nonetheless express a positivist concern: “Like causality itself, mechanisms are not directly observable. . . . How then are they to be imputed?” (p. 5). They suddenly exit the world of realism and climb back onto Hume’s stolid antirealist shoulders and announce they have solved the problem of the unobservability of mechanisms by invoking event uniformity: “Whenever the source of some event is unobserved, we should proceed on the hypothesis that it fits a pattern of causal uniformity. Causal uniformity implies the existence of a lawlike relationship that holds between events” (p. 6). The incoherence thus deepens: as Kiser and Hechter themselves have argued (pp. 4–10, 15–17), hypothesizing event uniformity (as in the covering law model) is *not* the same thing as finding causal mechanisms, just as demonstrating that childhood cancer is clustered in dense electrical power zones is not the same thing as demonstrating that (via direct causal sequences) concentrated electrical power *causes* cancer. Limiting an explanation to event uniformity is, after all, the only way to remain on empirical foundations.

Thus caught in the problem of trying to join the antirealism of the covering law model (which rejects the viability of causal mechanisms) to the realism of rational choice theory’s need for mechanisms, Kiser and Hechter’s only solution is to demand a strict inferential relationship in which causal mechanisms are “subsumed” under the causal law (1991, p. 6). Rather than solving the problem of how to find mechanisms, however, Kiser and Hechter appear to be simply waiving the problem away by retreating back into the safety of lawlike relations. Having first called mechanisms the “enduring focus of sociological explanations” (p. 15), they now want to conflate them into the uniformity of invariant laws. Struck by the awkwardness of this attempted merger of realism and antirealism, we might now ask: Whatever happened to the second requirement for adequate explanation—the causal mechanisms?

Theoretical Realism: Overcoming the Limits of Observation

Enter theoretical realism—the thesis that unobservable underlying realities can be ascertained by theoretical reason. Whereas the empirical re-

¹⁸ There are other reasons to object to the necessity of mechanisms; for the most powerful, see Stinchcombe (1991). What Kiser and Hechter really mean is that there is a consensus over the need for mechanisms among rational choice theorists (e.g., Cole-

quirements of positivism forced theorists from Hume to Popper to reject the search for deep causality beyond conjunction and probability, and whereas relational pragmatist realists use causal effects to infer theoretical entities, theoretical realism provides Kiser and Hechter a *deus ex machina* by making theory "see" what their senses cannot: with "general theory as the basis for the imputation of causal mechanisms . . . mechanisms must be imputed from general theories" (1991, pp. 5, 15, 16). Following the Lakatosian route out of Kuhn (through Popper) into the "hard core" of "general theory," theoretical realism thus allows Kiser and Hechter to accomplish by theoretical fiat that which has for centuries confounded philosophers—to bypass the epistemological problem of how to establish true causal explanation of mechanisms in the face of the problem of unobservability: "Models and mechanisms can only come from general theory" (p. 19).

What and Where Is General Theory?

We have arrived with anticipation at "general theory"—the fount of Kiser and Hechter's "adequate explanation." But, search as we might, it turns out that Kiser and Hechter never tell us what is general theory. To be sure, throughout they refer to it with a sprinkling of accolades designed to contrast with the weaknesses they impute to historical sociology—"theoretical boldness," "scope," "generality," "analytic power," and "parsimonious" (1991, pp. 9, 21). But the closest they ever come to a definition is to repeat that general theory is the "source" from which causal laws and mechanisms are to be derived. Indeed on page 17 we are presented not with a general definition but with a *single* case of a general theory—rational choice theory—and then informed that, to date, it is the *only* such case (p. 23). This presents a dilemma: The weaknesses, or for that matter the strengths, of rational choice—or any particular substantive general theory—is not the purported subject of Kiser and Hechter's argument, who instead give as their subject the *formal* criteria of methodology. Yet since discussion of these are missing at what turns out to be the crux of their entire argument, they have left us no choice but to use the particularities of rational choice to develop a "general" conception of general theory. We can thus add to Kiser and Hechter's difficulties: (1) by paying only the small price of an inconsistency in logic, they have smuggled in their preferred theory in the guise of general and neutral standards of epistemological adequacy, (2) they have thus violated the rules of epistemology and methodology, which, by definition, are required to define the general

man 1990; Elster 1989) The mistake is not an accident since they consistently use "theory" synonymously with rational choice theory.

criteria for sound knowledge and should not be derived from any one substantive theory, (3) and with this rather imperial move of conflating the specifics of rational choice with the definition of general theory across the board, they have closed off debate and preempted any possibility of methodological pluralism (Anderson 1993).

General Theory: Agents as Mechanisms

With these obstacles in mind, I begin by surmising what general theory is *not*. It is not in itself (although it may be used to impute) a set of testable propositions, either generalized from historical comparisons or deduced from empirical hypotheses. And it should not be (at the risk of tautology) either models or mechanisms, since Kiser and Hechter tell us those derive from general theory. In fact, in keeping with the hybrid character of their overall approach, Kiser and Hechter's general theory turns out to be a peculiar mix of aspirations—combining a methodological commitment to general laws (usually built on a staunchly empirical methodology) with a radically antiempirical theoretical ontology. Specifically, rational choice's general theory is a definitional bundle of assumptions about the nature of reality—a body of axioms that postulate the essential causal properties and powers of the social world. These essential properties and powers are mental states embodied in the *intentionality* of the individual agent: "All rational choice explanations begin with the assumptions that individuals are purposive and intentional actors who pursue prespecified goals . . . in rational choice theories, interests are specified a priori" (pp. 19, 21). Given metatheoretical form by Lakatos's (1970) influential theoretical realism, these assumptions are the hard core of the rational choice research program.

This conception of general theory is a classic case of what is called an "ontic methodology" (Salmon 1984); it is the *axial* core of all theoretical realism. The term is itself a hybrid: First it refers to an explicit *ontology*—a postulate about the nature of the social world. In keeping with theoretical realism's dismissal of the positivist prohibitions against nonempirical postulates in social science explanation, the ontic approach explicitly opposes positivism's privileging of epistemology over ontology—strict rules required to justify a knowledge claim ruling out any claim to theory and knowledge based on postulated theoretical entities. Theoretical realism thus begins with what positivism excludes: an empirically unknowable, and unfalsifiable, ontological theory of reality as a causal structure—an a priori argument about social reality based on "the idea of the essential causal powers of kinds of things" (Lloyd 1986, p. 8). In keeping with its antipositivism, ontic approaches are inherently driven by theoretical realism's fundamental distinction between "appearance" and "reality"—a dis-

tion as old as rationalism itself. Appearances are the mere "events" accessible to the senses of the observer; to mistake these for "reality" at the level of deep causal structure is to be empiricist. The theoretical realist, by contrast, privileges as the real source of causality the deep ontic structures and mechanisms of reality that are not accessible to the senses (Bhaskar 1979, 1986, 1989; Putnam 1997, pp. 181–84).

That rational choice has a core ontology is not in itself surprising; the protestations of some social scientists notwithstanding, all theories of knowledge make a more or less explicit ontological choice between either the individual or the social structure as the basic unit of social analysis (Alexander 1982). What makes Kiser and Hechter's approach different from a simple ontology is their combination of ontology with methodology—specifically, with rational choice's commitment to theorizing causal mechanisms. The result is not just a commitment to individualism, which could just as easily apply to interpretivist approaches. It is a commitment to a *causal* ontology in which agential intentionality is posited to be the a priori causal force/mechanism at work in the social world (Elster 1983, chaps. 1–3; 1989). An agent-centered ontology, to be sure, but more important is that as part of its very definition of reality it attributes exogenous causal power—via the property of mental intentionality—to its agential units of analysis. What completes the hybrid of an ontic methodology, then, is that a causal *explanation* (not just a postulated description) is not hypothesized but already built into the definition of social reality. This is a general theory—more precisely, an ontic explanatory presupposition less about agents as rational and more of *agents as mechanisms*.

Since all explanatory requirements flow from this presuppositional causal ontology, it is not rationality per se that predominates or even does the main explanatory work in rational choice's general theory; much more central is the causal capacity attributed a priori to the individual agents through the essential *disposition* of intentionality—which is endowed with inherent and exogenous causal power. Such an inherently causal ontology that postulates its agential units to have exogenous causal powers and generative forces makes use of what Hempel (1965, p. 472) called "broadly dispositional traits"—a term that evokes the ontological causal necessity characteristic of certain dispositions. The dispositional traits of intentionality and purposiveness are the essential but unobservable powers and properties I have stressed to be at the core of theoretical realism. As such, they fit the category of what philosophers call "reasons as causes" (Hempel 1962; Rosenberg 1988, pp. 27–30).¹⁹ Because they are dispositional, these mechanisms cause consequences with lawlike and self-

¹⁹ The locus classicus on "reasons as causes" is Hempel's "Rational Action" (1962). For discussion see Dray (1957), Rosenberg (1988), and Bohman (1991).

propelling necessity: It would be no more possible (all things being equal) for the agents of rational choice theory *not* to be driven by their intentional purposes than it would be possible for glass not to break on impact—given its essential dispositional trait as “brittle.”²⁰

At the heart of this general theory, then, is an ontic definition of reality composed of discrete agential units each endowed a priori with the dispositional causal power of intentionality. The explanatory work of the theory is carried by this invariant causal mechanism of a dispositional agential intentionality that necessarily (in the absence of constraint) causes intents to convert into actions be they rational or nonrational. Kiser and Hechter’s general theory has thus emerged to be a hybrid straining to make the center hold: It aspires to the model of the covering law, but it is a law in which the crucial explanatory causal mechanism is an a priori theoretical and ontological entity that is postulated exogenously. Rosenberg (1988, p. 23) calls this kind of law “folk psychology.” Rather than hypothesized, discovered, problematized, deduced, induced, or generalized, the explanatory content of this law comes ready-made; causality is inscribed exogenously. Independent variables are usually hypothesized rather than postulated, but in rational choice theory intentionality as causality is not only defined in advance but defined as the universal cause of all known effects. The causal mechanisms built into rational choice theory are always the same: Agents equipped with intentional dispositions that, universally and necessarily, are charged with a self-propelling momentum from cause/intentionality to effect/action. General theory, then, for Kiser and Hechter is a theory of agents as mechanisms believed to constitute and, by general universal law, to embody in their consciousness the exogenous and ontological causal power and force of intentionality that drives the social world.

From General Theory Back to Laws and Mechanisms

We have now reached the apex of a journey in search of how to produce “an adequate explanation” (Kiser and Hechter 1991, p. 4). It began with a call for laws (constant conjunction), argued it necessary also to find deep causes (causal mechanisms), and, having faced the problem of the unobservability of mechanisms, retreated back to laws and from there elevated to “general theory”—an ontology with a built-in explanation of agents as mechanisms. But what we have now learned is that they are also the very

²⁰ According to rational choice theory, there will of course be constraints and counterforces that serve as obstacles in any given empirical example, but they will be just that—external constraints on an otherwise natural dispositional force. Hence the *ceteris paribus* clause. The example of brittleness is found in Bohman (1991, p. 20).

same mechanisms that Kiser and Hechter tell us are the central component of all adequate explanation. What now? As a set of given assumptions, general theory can hardly be accountable to the petty trials of empirical confirmation or problematization; instead, now that we have found the source of the models and mechanisms in this general theory, we are about to start descending again: "Beginning with these basic assumptions, rational choice political sociologists then apply one or more of the causal mechanisms derived from available [rational choice] theories—such as power-dependence theory, repeated game theory, optimal location theory, agency theory, and group solidarity theory—to the problem at hand" (p. 19).

This is remarkable: The very causal mechanisms required for an adequate explanation of a problem at hand are to be inferred from a general theory composed of assumptions about *those same causal mechanisms*. Since the causal mechanism of intentionality is already inscribed—regardless of the specific case or the outcome being problematized—Kiser and Hechter predefine what should be an empirical question, namely, the question of causality in any given case. The tautology is inexorable: At the level of microtheory, because individuals are constituted in essence by the causal trait of intentionality, it would be impossible to explain any given action *without* attributing causality to their dispositional intentionality. And at the macrosocial level, likewise, it would be impossible to explain social outcomes without attributing causality to agential intentionality. With causal mechanisms *already inscribed* in the theory from which they derive these very mechanisms, Kiser and Hechter's ontic methodology gives new life to Plato's lament about the search for hidden truth: If we know what we are looking for, we have already found it; if we do not know, we cannot recognize it when we do (see Hollis and Smith 1991, p. 408).

Having It Both Ways

Kiser and Hechter have taken us on a dizzying circular trip: they have told us that something called general theory can give us the causal essence of a phenomenon, and from this essence they are handily able to explain a social problem by imputing to it the same causal property of dispositional intentionality that already lies at the definitional core of the general theory.²¹ The tautology lies in the absence of an independent explanatory

²¹ Tilly (1995c, p. 1595) spells out precisely the inexorable and tautological progression of this structure of explanation: "(1) assume a coherent, durable, self-propelling social unit; (2) attribute a general condition or process to that unit; (3) invoke or invent an invariant model of that condition or process; (4) explain the behavior of the unit on the basis of its conformity to that invariant model." This is Tilly's characterization

structure at least partially derived from the problem at hand. Because the explanation is already inscribed in the *a priori* definition of the problem, we are presented at the end with "little more . . . than the explication of a definition" (Tilly 1995c, p. 1595). Their commitment to laws notwithstanding, nothing could be more incompatible with standard positivist rules of deduction requiring testability and the capacity for falsification—or necessarily adjudicated by observation language.

When it comes to what counts as an adequate explanation, then, Kiser and Hechter try to have it both ways, in more ways than one. First, by combining the covering law model with the demand for causal mechanisms they aspire to the certainty of law with the elegance of theorizing deep causality. Then they try combining the whole mix with an axiomatic ontology. The problem is that the covering law model logic they rely on cannot be combined with theoretical realism's presumption that once a theory is determined to be true, the terms of the general theory denote the entities that are causally responsible for what we observe. The incompatibility derives from rational choice's theoretical realist definition of general theory as constituted by essential general properties, propensities, and powers—mechanisms, in short. By claiming to be able to explain by deriving both lawlike regularities/conjunctions as well as causal mechanisms from an axiomatic general theory, Kiser and Hechter are trying to combine positivism's antirealism with rational choice's ontic realism and in turn to deduce "how" mechanisms from this definitional theory—in advance of any empirical investigation—and not by engaging with the phenomenon itself. With this deeply anti-Humean/antipositivist solution in theoretical realism, Kiser and Hechter ultimately define general theory not by the Humean criteria they originally promised, but rather in anti-Humean realist terms as essential general properties, propensities, and powers. There is obviously no epistemological law that one must be Humean or positivist. Coherence, however, does require that since it is Kiser and Hechter's own choice to begin with Hume, and in turn to take on board the problematic of observability, they cannot claim to have found a coherent solution in what is a radically anti-Humean conclusion—a general theory composed of favored axioms about the imputed nature of reality.

CAUSALITY AND HISTORY

It would be churlish not to note the advance in Kiser and Hechter's desire for an adequate explanation to include mechanisms. In the covering law

of standard ways "to explain political processes"—not one that he applies specifically to theoretical realism. Yet it captures especially well the postpositivist deductivism of rational choice.

model, mechanisms are excluded not only because of problems of observability, but also because causal processes are considered too historical and particularistic, thus insufficiently logical and generalizable (Popper 1959; Salmon 1984; Bohman 1991; Miller 1987). Theoretical realism, by contrast, does not rest content with positivism's restrictive demands for observability as the criterion for explanation and insists, rightly, that cause cannot be ascertained by conjuncture alone; some understanding of mechanisms must be developed—how it is, for example, that the constant conjuncture of social revolutions on the one side and, on the other, a conjuncture of peasant rebellion and state breakdown, is actually a relationship of cause and effect? By insisting on mechanisms, then, Kiser and Hechter are to be commended for giving due importance to the processual and explanatory element of causality. Perhaps this is reason enough not to criticize the incompatibility of theoretical realism and classical positivism.

The apparent victory for the processual element, however, is Pyrrhic. Faced with the inability to observe causal processes, theoretical realism solves the problem by "observing" through the lens of theoretical assumptions the putative properties and propensities of unseen entities (e.g., the drive of human intentionality). No sooner do Kiser and Hechter turn to causal mechanisms than they declare that mechanisms can only be deduced from prior "omnitemporal, universal" laws, and then that both laws and mechanisms must find their source in a general theory whose analytic power lies precisely in its temporal abstraction (Kiser and Hechter 1991, pp. 6–7). This move allows their theoretical realism to reconcile its trafficking in metaphysics with the positivist goal of generating lawlike theory. The kinds of deductions about mechanisms that theoretical realism generates, given the definition of general theory as assumptions about essential properties internal to a contained entity, are deductions conforming to the same principles of omnitemporality, universality, and uniformity as those postulates of the general theory from whence they are derived. The result of this series of moves is an even stronger embrace than the original positivist one of the antitheoretical caricature of history.

For Kuhn, positivism wrongly neglected history; Kiser and Hechter take this temporal void to a level much higher than Kuhn ever envisioned. Whereas the standard covering law model conceptualizes time as "general linear reality" (Abbott 1988), Kiser and Hechter aspire to what can be seen as a nonlinear reality that does not so much linearize time as much as freeze it entirely. Theoretical realism, in fact, accomplishes a remarkable feat by solving the problem of causality not in the direction of Tilly's notion of "thick" causal time but in the direction of even greater invariance. Thus Kiser and Hechter "rule out historical narratives as a basis for the imputation of causal mechanisms," and instead demand "minimal

temporality" in selecting causal mechanisms (1991, pp. 6, 7).²² With this, theoretical realism in general and rational choice in particular have managed to resurrect what many had hoped would remain a long moribund false dichotomy between theory, science, and knowledge on the one side, and history, temporality, and narrative on the other. Surely it is ironic to find two historical sociologists telling us that the ideal explanatory structure is the one with the least temporality or "thinnest" conception of time—the very thing that has "repeatedly led social scientists to the mistaken conclusion that historians are particularizers while social scientists are generalizers" (Tilly 1994, p. 270).

HOW DO WE KNOW IF A THEORY IS TRUE?

Kiser and Hechter require an adequate explanation to be based on causal laws and mechanisms, but there is something missing in this definition—namely, a convincing reason for why we should believe in any given explanation. Explanations built on unobservable mechanisms with real causal powers operating with natural necessity require methods of confirmation very different from those built on observational correlations. But there is no trace of such an alternative approach to theory justification in Kiser and Hechter's argument. Positivists have always told us not to trust any theoretical claims that did not address the epistemological problem of how we can know whether a theory is true. Even if they were wrong in pointing to observation as the only foundation for claims to knowledge, they were nonetheless right to demand epistemological accountability. Kiser and Hechter's chain of inferences from general theory, to laws, to mechanisms may be logical, but in a post-Kuhnian universe logic is not enough to make a theory true. Kiser and Hechter still owe us an epistemology—a convincing account for why we should believe them.

In the same vein as their approach to an adequate explanation, the epistemology Kiser and Hechter do offer is cobbled together from a hybrid mix of pre-Kuhnian antirealist deductivism, and an antipositivist theoretical realism. The antirealist elements they use are: (a) a logic of justification limited to the dichotomy of induction versus deduction; (b) the uncritical acceptance of the superiority of Popper's version of deductivist logic. Their theory-centric realist elements are (c) the conflation of history, experiment, observation, and problem-driven research with "empiricism"; (d) an epistemology that tells us it is theory that adjudicates truth. The elements do not cohere: One cannot be committed to *b*, a standard positivist

²² Since that writing, rational choice theorists have discovered narrative—but only endorse a theory-driven version; see Bates et al. (1997), Kiser (1996).

deductivist logic of confirmation, in which truth is confirmed through observation language, and also adhere to *c*, a postpositivist theory-centered view of justification that builds on a theory-laden conception of data. Nor can *a*, a spectrum of justification limited to induction or deduction, be coherently joined with *d*, a theoretical realist epistemology in which truth is ultimately adjudicated by theoretical presuppositions.

The Limits of Pre-Kuhnian Deductivism

Kiser and Hechter lambaste what they allege to be the empiricist and inductivist practices of historical sociology (1991, pp. 5–9, 12–17). The philosophy of science, they claim, has ruled that inductivism is methodologically unacceptable (pp. 12–15), leaving deductivist logic as the only valid means by which to generate and confirm theory. Even within its own terms, however, Kiser and Hechter's argument is simply wrong. Deductivism does not hold this position of unequivocal triumph in the philosophy of science. To be sure, the theory-centric route out of Kuhn described above did, until very recently, dominate the history and philosophy of science and, superficially at least, seemed to converge with the primacy of deductivism (Hacking 1983, pp. 149–66; Miller 1987, pp. 453–61). But this has changed dramatically in recent years as revisionist demystifications of the textbook stories have converged with Kuhn's original criticism of both deduction and induction to undermine a disproportionately deductivist-driven view of scientific advance (see, e.g., Camic and Xie 1994).

A critique of deductivism however, also entails epistemological doubts independent of such recent corrections in the historical record. For one, Popper's criterion of falsification relies on "underspecified criteria" of what counts as decisive falsification criteria (Hull 1988; Bohman 1991; Hollis 1994; Miller 1987), just as it neglects the fact that theoretical hypotheses often make no claims about observation directly and thus cannot be tested except when supplemented with other hypotheses. This allows an escape clause for the hypothesis of choice by simply attaching blame for a recalcitrant observation to flawed measuring devices or to the faultiness of the supplemental hypothesis, thereby leaving the favored one intact (Quine 1963, 1969; Hollis 1994, pp. 79–80; Rosenberg 1988, p. 47). It is obvious in the social sciences, moreover, that the test of falsification—in which a single counterobservation can falsify a theory—is virtually never practiced (consider Marx's theory of class formation; see Somers 1996a); and for good reason, given that more than one theoretical construction can almost always be placed on a body of evidence (Kuhn 1970, p. 76).

Equally important is the frequency with which well-established scientific theories survive and thrive despite bearing the burden of rather

mind-boggling implications: Until the last few decades, for example, the theory that matter is made of molecules, when combined with the best-established theories of force, implied that no floor could hold a person's weight. Deductivism fails here because it contains no criteria for distinguishing between respectable reasons for defending a theory against such anomalies and dubious tactics—for example, dismissing anomalies based on insufficiently similar conditions, yet without specifying the criteria for sufficient similarity (Miller 1987, p. 232). Deductivism thus provides no way to prevent dogma from thriving quite comfortably under the guise of science.

The obverse side of the stubborn ability of most hypotheses to survive the test of falsification is the even more privileged—and problematic—status of corroboration accorded to those theories for which disconfirming data has not *yet* been found. The theoretical indeterminacy of data, however, makes suspect the idea that there will ever be firm consensus on what data counts as disconfirming in the first place—thus setting almost no limits to the criteria for corroboration. Even more hubristic are claims to successful corroboration based on what is, as of yet, merely an absence of counterevidence. On this principle, the more abstract the theory and the less amenable to empirical falsification, the more likely it is to be considered valid. That planets of stars other than the sun are Camembert moons, for example, could be considered corroborated in the absence of any actual observations to the contrary. Through failure to falsify, “All emeralds are green” and “All emeralds are grue” have both been corroborated (Miller 1987, pp. 235–36), just as failure to yet falsify “all post-Soviet capitalist regimes lead to socialism” allows it to stand as corroborated.

What these problems point to is the difficulty of proving or disproving a theory on purely logical grounds—the essential core of the deductivist claim to scientific superiority. Yet this is precisely what Kiser and Hechter do. Skocpol (1994, p. 323), for example, rightly points out that Kiser and Hechter attack her work “not by dealing with *any* of its substantive ideas or findings, but by showing—in very general logical terms—that it does not live up to what they claim, in the abstract, any ‘adequate explanation must entail.’” They likewise dismiss on logical grounds a vast array of substantive findings in historical sociology—ranging from world economies to contentious movements—on the grounds that these findings have been produced through the failed logics of “empiricism” and “induction” (Kiser and Hechter 1991).

In one (by now well-worn) such example of a logical rather than a substantive criticism, Kiser and Hechter attack historical sociologists for their lack of ability to achieve case independence (p. 13). In addition to ignoring the work in network theory that takes issue with the logic of this criticism, their attack also presupposes that cases are discovered as ready made,

discrete entities—given in the nature of things, that is—rather than constructed along analytic parameters. In problem-driven research cases are not selected from an external spectrum of possibilities; cases are *defined* by the *problem* the researcher has set out to explain. Historical sociologists, by and large, thus draw the lines of independence between cases based upon what it is they are trying to explain; case selection is an analytic, not a logical, operation. The goal of the research design is to establish that an object is an adequate empirical unity for the limited purpose of the problem at hand.²² Thus Skocpol reminds us that she “could and did make comparisons” between cases she carved out as temporal “episodes” of sociopolitical conflict within single countries because she did not passively assume she could only compare discrete and predefined “countries” (1994, p. 322). Rather, she actively defined her cases as analytic temporal episodes precisely so that she could test her problem-driven hypothesis about the impact of Russia’s state apparatus as it weakened over time (1905–17). Similarly, both Putnam (1993) and I (Somers 1993) draw theoretical conclusions by comparing contrasting political cultures across different regions of a single country—cases that would be ruled nonindependent (because in the same national society) according to an abstract logic that leaves unproblematicized the crucial question, What is a case? (Ragin and Becker 1992). In all of these examples, arguments from abstract logic overlook the valid findings that emerge when cases are configured according not to the abstract logic of deductivism, but to “the logic of the causal hypotheses being presented and tested” (Skocpol and Somers 1980, p. 194). Independence is defined analytically according to what it is that is being problematized and thus along substantive, not logical, dimensions. Following Marc Bloch’s (1934, p. 81) maxim about the comparative method, “Only the unity of problem makes a center.”

Postpositivist Deductivism

Kiser and Hechter specify three criteria to adjudicate among competing theories: “plausibility, reduction of time lags between cause and effect, and the empirical implications” (1991, p. 6). Despite appearances, each of these criteria, including the third, is premised neither on logical nor empirical grounds but on the theoretical ones of rational choice theory, thus suggesting that, for Kiser and Hechter, the test of a good theory is whether it meets the criteria of the particular theory favored by the researcher.

²² What counts as causal unity is an empirical question and so can change over the course of research. Very often one starts with commonsense causal unities—e.g., people have intentions—that with additional information can be transformed or extended to less commonsensical ones—e.g., firms have intentions.

This represents a hybrid of antipositivist theoretical realist deductivism the unworkability of which is manifested in the circular reasoning they employ in discussing these criteria. First there are criteria based on logic: "The success of such studies ultimately rests on the degree to which they meet the requirements of good causal explanations" (Hechter 1992, p. 368). Then, in light of the recognition that one of their logical criteria—providing causal mechanisms—relies on unobservables they turn to theoretical realism: "We cannot go out and collect the data [on causal mechanisms]. . . . On the contrary, this is what we need general theories for" (Hechter 1992, p. 368). This logic generates inferences that tautologically confirm their hypothesized ontology of the world, thus allowing theory to provide the data by which that same theory is judged.

Consider next the criterion for theoretical success being the greatest reduction of time lags between cause and effect.²⁴ Citing Elster (1983, p. 24) Kiser and Hechter state that the best way to reduce the time lag between cause and effect is to replace macrovariables by microvariables or to link macrolevel variables using intervening microlevel ones (p. 7). But in this context microfoundations are not logical criteria; they are the central metatheoretical principles of rational choice (see, e.g., Elster 1989). It thus follows inexorably that their criticisms of historical sociology (p. 16)—historical sociology's "inadequate microfoundations"—turn out to be neither logical nor empirical, but disagreements among rival *theoretical* views of the world. Kiser and Hechter are telling us, unabashedly, that the presence of microfoundations must be the *general* epistemological criterion for the success of a theory. In this instance, mine is not a complaint against rational choice; it is an complaint against an epistemology that adjudicates the success of any given theory by axioms derived from the very essence of that theory.

With their third criterion of testability, Kiser and Hechter now appear to follow Friedman's (1953) "positive economics" in which axiomatic assumptions are just that—axiomatic "as if" heuristics, not meant to be realistic but designed to meet the instrumental test of producing testable predictions (see also Bell 1980). From this perspective it matters not that hypotheses are inspired by unrealistic assumptions as long as they are kept logically distinct from the methods of justification by which the hypothesized predictions are to be tested—thus falling back on Popper's distinction between discovery and justification. But Kiser and Hechter

²⁴ Putnam (1993), not unsympathetic to rational choice, would disagree. After extensive testing he found the 14th century to be critical for his explanation of civic democracies, just as I found, after testing three competing hypotheses, that England's 12th-century legal foundations were central to citizenship formation (Somers 1994b). And for another very long-run causal analysis see Tilly (1990).

are also postpositivist theoretical realists. Realists, following Kuhn, deny this distinction on the premise that in deciding which facts to select and transform into evidence we are already to a large degree deciding among rival theories (Hollis 1994, p. 79). Kiser and Hechter, moreover, make it clear that their use of general theory is much more than heuristic; it is the source of the causal mechanisms they will use to describe and explain the world. To concede their general theory as no more than heuristic or instrumental (in the Friedmanite sense) would be to concede to positivism's antirealist foundational premise that true causal knowledge about the deep structures of reality is impossible to establish. With this concession would collapse the entire rational choice project.

This dilemma is not unique to Kiser and Hechter but characteristic of rational choice as a whole. When rational choice theorists justify the "as if" status of their assumptions on the grounds of instrumental utility in generating predictions, they inevitably contradict their simultaneous demands for theory structured by either general covering laws or by (true) causal mechanisms. Referring to the first incompatibility, Green and Shapiro point out that "either the development of general theory is justified on covering-law grounds (in which case it cannot legitimately be based on unrealistic assumptions), or the unrealism is justified on instrumental grounds (in which case the particular mode of theory building is beside the point; testable predictions are what matter)" (1994, p. 31). And, with respect to the second, one cannot claim that theory is both the grounds for establishing a description of reality on the one hand and, on the other, a mere analytic language of heuristics for which the criteria of realism is unnecessary to meet its prediction-generating goals. As Friedman would be the first to concede, causal mechanisms telling us *how* something happened cannot be deduced from "as if" assumptions designed to produce only empirical regularities (Friedman 1953).

On close examination, even Kiser and Hechter's criterion of testing via "empirical implications" turns out to be more theoretical than empirical as they appear to test the power of a theory to predict something the theory is *already* premised upon—the free rider problem is one such example (pp. 8–10). The limits of these overly theoretical criteria cannot be overcome by the claim that rational choice predictions test "retroductively" (Fiorina and Schepsle 1982, p. 63). One reason is that in the social sciences when theories are judged retrospectively, more than one almost always will be interpreted to predict consistently with the facts. Another is that the kind of consistency Kiser and Hechter claim is not one in which an empirical puzzle is explained by a theoretical model that fits both positive and negative cases of the same phenomenon. Rather the kind of prediction and testing that Kiser and Hechter use to exemplify the superiority of rational choice's theory-driven approach over the problem-driven one of

most historical sociologists, is a post hoc one: Facts that are already known are explained in an account fashioned to be consistent with rational choice's a priori postulates about essential (thus ultimately untestable) faculties and properties of theoretical entities such as intentionality. Such post hoc explanations that are consistent with a theory's original assumptions amount to little more than the ability to rewrite in a "tribal language" (Skocpol 1994, p. 325) what others have laboriously discovered. Moreover, given the "lack of specificity about what it means to be a rational actor, it is not obvious what sorts of behaviors, in principle, could *fail* to be explained by some variant of rational choice theory" (Green and Shapiro 1994, p. 34; emphasis added). Telling a rational choice story about data that has already been observed hardly qualifies as passing the testability that Kiser and Hechter repeatedly demand of adequate theorizing. "Data that inspire a theory" we are reminded by Green and Shapiro (1994), "cannot . . . properly be used to test it, particularly when many post hoc accounts furnish the same prediction. Unless a given retroductive account is used to generate hypotheses that survive when tested against other phenomena, little of empirical significance has been established" (pp. 35–36).

Finally, Kiser and Hechter use yet another set of timeworn criteria for good theory—"parsimony," "universality," "scope," and "boldness." But these fare no better than the testability claims. Why should we prefer what are, after all, aesthetic qualities to, say, the "complexity, expansiveness, and historicity" advocated by Dessler (1989, p. 446)²⁵ or by Hirschman (1984) in his classic statement, "Against Parsimony"? The only convincing reason, as Hollis (1994) reminds us in referring to Friedman's (1953) similar demand for parsimony, is that they believe these qualities actually capture the *reality* of the world being theorized—that is, the world really *is* parsimonious and invariant. One could take issue with this ontology, as indeed I do below, but certainly one can neither confirm nor deny it—thus making it a justification wholly at odds with Kiser and Hechter's original premise that the ultimate test of a theory is its empirical testability.

To sum up: In their criteria for theory confirmation, Kiser and Hechter splice their postpositivist theory-centrism together with the positivist language of deduction. Their assumption that the only available alternatives in theory construction are the inductive/deductive dichotomy at first glance appears to make them merely strong positivists on the deductivist side of the debate. But standard deductivist logic confirms/tests a theory by *observing* the world—so much so that, as I have emphasized, it dismisses the possibility of justifying knowledge claims in the absence of

²⁵ "The richer and more comprehensive the underlying ontology, the better the theory" (Dessler 1989, p. 446).

observable evidence. Kiser and Hechter's version of deductivism, by contrast, is the aggressively antipositivist deductivism of theoretical realism. It does not fit Popper's requirements for theory testing by data ("conjectures and refutations") for it is a deductivism in which hypotheses are not only guided but *adjudicated* on the basis of general theory. A post-Kuhnian theoretical realism combined with a pre-Kuhnian deductivism thus creates a hybrid suspended between, on the one side, what Kiser and Hechter call "empiricism," and, on the other, a theory centrism removed from the need to observe the world.

We have now come full circle. General theory is both the judge and the jury in an epistemology that, to paraphrase Hollis's (1994, p. 36) remarks on rationalism more generally, makes the purpose of a theory (of action, e.g.) to find the essence of any given phenomenon (a given action) by defining that concept (of action) in a way that (inevitably) captures that essence. In the pattern of this rationalist epistemology, Kiser and Hechter dismiss what they define as historical sociologists' overly great concern for accuracy and empirical detail as "naive empiricism" when they compare it against the standard of their disproportionately deductivist view of theory. Here is an epistemology truly grounded in the assumption that the test of a good theory is its ability to trump observation on terms set by the *same* theory. A theory is thus to be admired when it "outpaces" the data that is to be placed in its pigeon holes (Lakatos 1970). The problem remains: We still have been given no good reasons to believe in it.

THEORETICAL REALISM: AN ONTOLOGY FOR THE ANGELS

In his recent discussion of how to explain political processes, Tilly (1995c) criticizes sociologists' widespread use of "invariant models concerning self-motivating social units" (p. 1596). His criticism is important and unusual; in lieu of the common exclusion of ontology ("the nature of that which is to be known") from discussions of social science methodology, Tilly insists that because ontological assumptions inevitably influence epistemology ("the conditions for the generation of knowledge"), it is crucial that those assumptions be "plausible" (p. 1602).²⁶ In keeping with this approach and picking up on my discussion of rational choice's "ontic methodology" as well as its use of "reasons as causes" and "agents as mechanisms" (see pp. 749–51 above), let us examine first how Kiser and

²⁶ Tilly is hardly alone in recognizing the centrality of ontological assumptions on methodology. See, e.g., Alexander (1982), Giddens (1984), and all of the literature on realism (e.g., Bhaskar 1979, 1986, 1989; Lloyd 1986, 1993).

Hechter relate their epistemology to their ontology, and in turn take seriously their own suggestion (1991, p. 6) to consider the "plausibility" of the ontology they do embrace.

In the first instance, rather than making their form of explanation dependent upon the character of the entities whose existence they are prepared to defend, Kiser and Hechter derive their ontology from their prior claims about what logical forms of explanation are acceptable. This strictly positivist approach reflects a three-tiered method that exhibits the following logic: (1) ontological claims ("X exists," "X has these properties") are licensed *only* once they have been validated by a general theory of X; (2) we can capture the true intrinsic character of X (i.e., the properties, including causal powers, that X possesses) only through causal relations that relate X's to Y's to Z's properties according to the covering law model; (3) since we already know X's ontological properties through the postulated causal assumptions applicable to any given X, it is not necessary to examine empirically the causal mechanisms through which Xs are related to Ys. (In rational choice theory, recall, causal mechanisms are held to be derivative from and entailed by X's a priori ontic postulates of intentionality.)

At odds with positivism's privileging of logic with which they started, Kiser and Hechter thus ultimately ground their logic on a realist premise—that laws and mechanisms must be imputed from general theory. Since, as I stressed above (pp. 749–52), their general theory turns out to be an exogenous ontology of the causal structure of reality, in this second instance they are donning their antipositivist theoretical realist hats and rejecting the priority of logic by putting their ontology first. If, as Tilly suggests, this influence of a prior ontology on one's epistemology is to some degree inevitable, then indeed we must ask, How "plausible" is their ontology?

The basic elements of rational choice are well known—that actors apply the standards of means-ends rationality, that they are self-interested, and they are wealth maximizers. These claims have been challenged forcefully at every turn and from every conceivable angle, and it is not my purpose to repeat those challenges.²⁷ However, as I have stressed throughout, these substantive elements are not the foundational core of the theory. More important is that which is at the heart of Kiser and Hechter's ontic methodology. Recall the basic agential units of analysis in rational choice: Not just agents, *qua* actors, but agents that come already

²⁷ Among the most important are Anderson (1993), Block (1991), Hollis (1995), Green and Shapiro (1994), Plizzorno (1986, 1991, 1995), and White (1992). For a recent critique applied to the history of state formation see Adams (*in press*).

equipped with the essential properties of purposiveness and intentionality that will *by dispositional necessity* (Hempel 1965; and see above pp. 750–51) cause them to “pursue prespecified goals” (Kiser and Hechter 1991, p. 19). What emerges, then, is that the basic agential units are *in themselves* causal forces—what Tilly (1995*b*, p. 1595) calls “self-propelling” entities, and what I have called above “agents as mechanisms.” The theory that posits its causal mechanisms to be exogenously borne by human agents is also necessarily an ontology that assumes its agents to be inherently self-galvanized by self-contained autonomous mental states (reasons)—thus making them “coherent, durable, selfpropelling social units—monads” (p. 1602).

But it was no less that Karl Popper (1959) who famously took issue with what he called “essentialism”—a philosophy which looks to the “essence” of things for information about their “true” nature and behavior. An essentialist approach treats its objects of analysis as constituted by a set of inherent attributes—attributes intended to represent the essence of a thing. For Kiser and Hechter and rational choice, this essence is the exogenous causal power of intentionality. Essentialist ontologies, in this sense, are recognizably presocial, almost Hobbesian (Pizzorno 1986, 1991). They posit not only fixed solipsistic identities but ontological entities that are born preequipped to act through essential inherent causal mechanisms (reasons as causes) that drive action on their own autonomous momentum. Questions we would want to ask about this ontology include: How we can know or justify knowledge claims about the essence of unobservable mental states in the first place? Can the single putative property of intentionality define a human being, let alone explain or predict human action? How is it possible to claim *social* agency for an identity if its motivating force derives from apparently presocial or fixed categories constructed from exogenous attributes (Somers and Gibson 1994, p. 55)? And is it not possible that these “essential” identities are, as White (1992, p. 8) points out, “by-products of previous history adapted to current circumstance . . . not causes, but rather . . . spun after the fact as part of accounting for what has already happened”?

Thus postulated on a dispositional ontology of theoretical entities (states of mind) with a priori causal mechanisms (the intentionality of agents as mechanisms), I find implausible Kiser and Hechter’s perception of the social world—namely, that it really is composed of agents with essential and unchanging properties that operate independently of the very relationships by which they are constituted. The implausible implications of an ontology in which causality is dispositional, rather than investigated, point to a world of invariance, a “world in which whole structures and sequences repeat themselves time after time in essentially the same form”

(Tilly 1995c, p. 1602). Kiser and Hechter's ontology of agents-as-mechanisms invokes a world populated by "angels"—ontological entities that, implicitly, are not of this social world. Invariantly constituted by the singular dispositional drive of attributes and essential intentionality, these agents are defined as having discrete and autonomous states of mind constituted by causal mechanisms that are not only given a priori by theoretical assumptions, but inexorably do their work in isolation from other entities or processes. No wonder Kiser and Hechter are so confident of their predictive capacity: A world of angels possesses few surprises, founded as it is on invariant properties and self-propelling causal powers that are knowable a priori. To be sure, as Tilly ponders, "If the social world actually fell into neatly recurrent structures and processes . . . invariant model and the testing of deductive hypotheses *would* become more parsimonious and effective means of generating knowledge" (1995c, p. 1602; emphasis added). The problem, however, is that while this "would be a convenient world for theorists . . . it does not exist" (p. 1596).²⁵

Rational choice theorists seem to disagree. As theoretical realists they presume the possibility of "real and relatively atemporal" objects of analysis that lend themselves to "objective and progressively successful discoveries and explanations of them" (Lloyd 1986, p. 9). For theoretical realists this conception of the categorically inherent properties of reality is the final adjudicator of reason. This assumption generates logical inferences that tautologically confirm their hypothesized ontology of the world—its universal, omnitemporal nature.

Kiser and Hechter's theoretical realist ontology is at root a metaphysics—an a priori belief that the social world really can be theorized as comprised of universal, invariant, entities with discrete self-propelling causal dispositions and agents as causal mechanisms. There is an irony in this: As realists, Kiser and Hechter very much believe that a real, mind-independent world exists "out there." Yet, absent the relational, contingent, and causally indeterminate element of thick time and being, it becomes a thin world of pure idealized spirit (Stinchcombe 1978, p. 21). At the end of the day, one can only agree with White that Kiser and Hechter's rational choice vision rests on "an underlying ontology of 'spirits' . . . upon angels, that is, upon spirits both disembodied and independent. . . . [The] goals or preference orderings . . . essential to . . . rational choice, are appropriate and relevant only to entities which are inertial as well as isolate—

²⁵ Hence the great success of Green and Shapiro (1994) is to show that rational choice has failed empirically in at least three of the most important areas it has theorized in political science. See also Friedman (1996) for a range of such critical perspectives.

angels, in short" (White 1992, p. 301). A world of angels may be parsimonious and convenient to theorize, but—disappointing, to be sure—we're no angels.

RELATIONAL REALISM

Relational and pragmatic realism is a post-Kuhnian perspective for the rest of us—those who are not angels. It should be understood as a minimalist realism—minimalist because it recognizes that the partial concept-dependence of social life puts limits on the general realist premise of the absolute mind-independent status of the social world; yet realist nonetheless, in contrast to hermeneutics or radical constructivism in that some degree of concept-dependence does not in any way subvert the premise of a social world that exists independently of our beliefs about it. Relational realism thus has three limiting principles: First, belief in the causal power of a theoretical social dynamic (e.g., gender, utility maximization, class struggle), is independent from belief in any *one* particular theory. Following Kuhn's (1970, pp. 173–74) mandate to think of theory advance as movement *from* present knowledge, rather than *toward* absolute truth, this injunction endorses Taylor's (1989) notion of "epistemic gain" in which knowledge is understood to be limited to "movement from a problematic position to a more adequate one within a field of available alternatives" (see also Calhoun 1995, p. 36). Second, all theory, especially one like rational choice that is premised on an a priori causal ontology, must provide an epistemology—some very good reasons—for why we should believe it (Hollis 1994; Hollis and Smith 1991, 1994). And third, there are no universally valid principles of logical reasoning; there are only problem-driven ones (Kuhn 1970; Lloyd 1993; Miller 1987, p. 486). Since space prohibits me from detailing the complexities of relational pragmatic realism, I merely sketch out and signal the markers of this route that has begun to bear out the potential challenge of a historical epistemology.

A Relational Realist and Pragmatist Ontology

A relational realist and pragmatist ontology is for those of us who accept, however unwillingly, the brutal fact that we and our social world are not angelic, existing outside time and space, but living, breathing, changing, dying creatures and entities, embedded in time and constituted—not merely engaged—in relationships. Beginning with the postulate that we are neither monads nor self-propelling entities but "contingent, transitory connections among socially constructed identities" (Tilly 1995c, p. 1595), a relational pragmatist ontology takes the basic units of social analysis to be neither individual entities (agent, actor, person, firm) nor structural

wholes (society, order, social structure) but the relational processes of interaction between and among identities (Collins 1981; Pizzorno 1986, 1991, 1995; Stinchcombe 1991, 1992, 1993, 1995; Tilly 1995a, 1995b, 1995c). Insofar as professional historians have developed ways to represent these processual and relational characteristics of "the nature of that to be known," relational realists borrow freely from what has traditionally been historians' terrain—through the appropriation of narratives, for example, as a way to represent sequences and processes over time (Abbott 1992a, 1992b, 1993, 1995; Aminzade 1992; Griffin 1993; Somers 1994b, 1996b). They do this not because they embrace historians' *norms* as Kiser and Hechter accuse; if they choose narrative *forms* it is because they hold the ontological belief that the world is made up of things that are constituted through temporal and spatial relationships and thus *must* be represented in relational and narrative terms (Wallerstein 1991). Hence, for example, agency can be usefully represented through the concept of a *narrative identity* to signify it as "processual and relational . . . [it] embeds identities in time and spatial relationships," and analyzes all identities "in the context of relational matrices because they do not 'exist' outside of those matrices" (Somers and Gibson 1994, pp. 61, 65).

A relational ontology thus follows Popper's rejection of essentialism and instead looks at the basic units of the social world as relational identities constituted in relational configurations. In place of a language of essences and inherent causal properties, a relational realism substitutes a language of networks and relationships that are not predetermined but made the indeterminate objects of investigation. Relational subjects are not related to each other in the weak sense of being only empirically contiguous; they are ontologically related such that an identity can only be deciphered by virtue of its "place" in relationship to other identities in its web. What appear to rational choice theorists to be autonomous agents defined by their attributes are thus better reconceived as historically shifting sets of agential relationships contingently stabilized in sites.

One need not be a historical sociologist to embrace such a relational ontology; it entails no more than the premise that identities, and hence potential mechanisms, are constituted—not merely constrained—in variable relationships. Following Adams (in press), it does not require a rejection of rationality but simply treating the "disposition to act rationally as a variable rather than a postulate" (p. 26). Thus, as per Smelser (1992), it requires us to conduct research into "the question of the contextual conditions—motivational, informational, and institutional—under which maximization and rational calculation manifest themselves in 'pure' form, under which they assume different forms, and under which they break down" (p. 404). And of course what is true of the variability of rationality is equally true of relationality: "Relationships may be more or less bonded,

the experience of them may be more or less constricting or enabling . . . but this is a question of contingency . . . [for] relationality is an analytic *variable* instead of an ideal-type" (Somers and Gibson 1994, pp. 61, 65; emphasis in original).²⁹

Relational Realism and Adequate Explanation

What counts as an adequate explanation for relational realism can be distinguished in several ways from theoretical realism in its convergence with elements of a post-Kuhnian historical epistemology. First, the basic mechanisms of causality are not within discrete agents, but in the *pathways of agential interaction* or what Stinchcombe (1991, 1992, 1993, 1995) calls and demonstrates to be contingent and empirically variable "situational mechanisms." Relational realism thus takes on board rational choice's criticism of functionalism's neglect of mechanisms without reverting to the ontic methodology of what I have been calling agents-as-mechanisms. That causal mechanisms are in contingent relational pathways rather than in the minds of self-propelling agents is supported by the social sciences' increasing recognition of the necessity of explaining outcomes through path dependency and sequence analysis (Abbott 1992*b*, 1995; Aminzade 1992; Arthur 1988, 1989; David 1985; Aminzade 1992; Tilly 1995*a*, 1995*c*; Putnam 1993; Sewell 1996; Stark 1992), and causal narrativity (Somers 1994*b*, 1996*b*). Adumbrated in Kuhn, and brilliantly demonstrated in David's (1985) renowned economic analysis of the QWERTY keyboard, path dependence and causal narrativity incorporate a sense of thick causal time in which "for any given trajectory, past choices and temporally remote events can help to explain subsequent paths of development and contemporary outcomes" (Aminzade 1992, p. 462; see also Skocpol 1979). When incorporated with comparative trajectories (as is ideal) these methods allow us to hypothesize, theorize, and test the causal status of previously embedded institutional practices; 14th-century legal institutions, for ex-

²⁹ Rejecting the either/or between individualism and holism, between agency and structure—but without resorting to the "mutuality" compromise—relational realism thus finds an analytic home in network theory (Bearman 1993; Breiger 1991; Burt 1992; Laumann, Marsden, and Prensky 1983; Padgett and Ansell 1993; Wellman 1988; White, Boorman, and Breiger 1976; White 1992, 1993, 1996), narrative/sequence analysis (Abbott 1992*a*, 1992*b*, 1995; Aminzade 1992; Griffin 1992, 1993; Somers 1994*b*, 1996*b*, Steinmetz 1992; Stinchcombe 1978; Tilly 1995*a*, 1995*b*, 1995*c*), space-time structures (Aminzade 1992; Sewell 1996; Wallerstein 1991), relational identities (Pizzorno 1991, 1986; Somers 1992, 1994*a*; Somers and Gibson 1994; Tilly 1984, 1995*b*; White 1992), configurations (Skocpol 1984; Tilly 1984, 1988), and historical institutionalism (Steinmo, Thelen, and Longstreth 1992; Somers 1993), each of which fits loosely under the general rubric of what Emirbayer (1997) has recently declared to be a burgeoning *relational sociology*.

ample, can be shown under certain initial and subsequent conditions to neither disappear nor become "lag effects" with the onset of capitalist markets in the 18th century, but to be causal factors in the development of 19th-century democratic institutions (Somers 1994b; see also Putnam 1993). Path dependence and causal narrativity suggest that earlier institutional processes are "sedimented" into the core of some of our most modern phenomena; they thus help to explain why some aspects of the social world, and not others, became modern in the first place. One of the most significant contributions of these processual and relational methodologies is that it makes it possible to theorize not just social change (long the sociologist's fixation), but the equally important phenomena of social durability and institutional patterns of resistance (e.g., the QWERTY keyboard, or the English common law). And although no one has done more than Abbott to demonstrate how an explanation that depicts mechanisms is best constructed through narrative sequential structures, it is telling that it was Mann (1986), Skocpol (1979), Stinchcombe (1978), and Tilly (1984, 1990)—four of Kiser and Hechter's core targets—who were among the earliest to demonstrate empirically and historically that outcomes could only be fully accounted for through causal explanations that incorporated degrees of path dependency.

The second aspect of a relational realist explanation decouples the two aspects of cause that form the centerpiece of Kiser and Hechter's argument—laws and mechanisms. Relational realism insists that "how" questions of causality (processes/mechanisms) cannot be deduced either from invariant laws (constant conjunctions/relations) or from the exogenous essential causal properties and powers that constitute Kiser and Hechter's notion of general theory.³⁰ Relational realists justify uncoupling mechanisms from laws by (a) the contingent and indeterminate nature of the basic causal mechanism (the relational unit of interaction) and (b) the fact that the adequacy of an explanatory structure depends upon why a question is asked: Different research aims yield different explanatory demands. Laws may be feasible for prediction, but only causal mechanisms accounting for variation and relational linkages can explain *how* and *why* something actually happened. Neither fundamental laws nor imputed self-propelling causal properties and dispositions govern objects in time and space; they only govern objects in abstraction (Cartwright 1983). The answer to a "how" question instead demands explanatory and empirically contingent causal pathways. Hence there are the respective rules for the two types of accounts: Competing theoretical treatments—treatments that offer different general laws for the same phenomena—are encouraged in

³⁰ Paradoxically, that mechanisms cannot be deduced from laws is also the argument of Elster (1989).

physics; by contrast, only a single causal story is allowed. Alternative causal stories compete in physics in a way in which covering law treatments do not. Causal stories are treated as if they are true or false, but which fundamental theories govern the phenomena is a matter of practicality (e.g., there are dozens of fundamental theorems for laser operations and scientists openly choose one or another depending on other factors; Cartwright 1983, pp. 11–12). And not surprisingly, although philosophers generally believe in laws and deny causes, the specificity of causal stories means that in actual practice physics works in just the reverse.

Let us take a classic example from economic theory. The fundamental law is based on microeconomic theory and it purports to explain how firms make decisions about prices: Because they are motivated to maximize profits, managers determine prices by setting output at a level where marginal costs and marginal revenues are equal. This may work as a powerful predictor about a relationship between managers and prices but it certainly does not give us an explanation for how prices are set. Why not? Because it is undisputed in economics that no manager or any economist has the slightest idea of what the marginal cost of producing something really is. For a cause to explain, the cause really has to exist; it has to be identified and exhibited. An “as if” underlying theoretical postulate is *not* a cause, it exhibits no evidence. No less than Milton Friedman has admitted that the actual thought processes of managers cannot in practice resemble this model. We have a covering law, but we do not have an explanation.

Since causality, unlike general laws, involves “stories” constituted of causal mechanisms and connections, the thick temporality of path dependence and causal narrativity makes them especially significant in the defense of history against the attacks of Kiser and Hechter. Take an example with which we are all familiar: It is a general covering law that to lose weight, you have to eat less. This *suggests* causal mechanisms—eating less causes the effect of losing weight—and it may also be a powerful predictor, but it is a correlation of events that does not in fact demonstrate the mechanisms that actually can explain *how* it happens. The law does not exhibit a cause. How do we exhibit a cause? We tell a story, a causal narrative about the causal pathways by which one class of events is actually affected by another. Since food has calories, and calories are energy, when we reduce our intake then the body has less energy to draw from external sources so it has to turn to internal sources of energy, which are stores of fat, and the body uses up fat when it draws that energy, and so forth. Along the way we may use general laws but they are not in themselves explanations for why and how, when we eat less, we lose weight. In science as well as sociology, an explanation that actually depicts causal mechanisms is always told in narrative form comprising statements with

transitive verbs: "The reduction of energy caused the body to draw on other sources . . ." "A actually caused B to happen by means of the following mechanisms and processes. . . ." Cause implies narrative and relational pathways. It is narrative because explanation must be embedded in time and move through time. Indeed the success of any explanation resides in its accounting for temporality and sequence.

Contra Kiser and Hechter, a relational realist explanation thus entails no trade-off between the false dichotomy of theory versus history. Following Kuhn's mandate to liberate history from its long-demeaned status in epistemology, relational realism is devoted fully to generating causal explanations but from an ontology based on relational mechanisms. At issue in proposing a causal theory based on causal narrativity and path dependency is *not* (as Kiser and Hechter insist) the choice to eschew causality and theory in favor of particularism; rather it is the choice to *generate causal theories that build on the temporal processes that are at the heart of any notion of a causal mechanism* (e.g., Coleman 1986, 1990). Relational realist methodologies demonstrate that when causal mechanisms are no longer "subsumed" (Kiser and Hechter 1991, p. 6) covering laws, we are left not with a naive particularism, as Kiser and Hechter suggest, but with Tilly's (1994) notion of time as one "drenched with causes that inhere in sequence, accumulation, contingency, and proximity" (p. 270).

Path dependence and causal narrativity converge with recent work in biology that suggests that many of our natural laws may hold *because* they can explain things by reference to a common history (Hull 1977, 1978, 1981). Comparing historical sociology to evolutionary biology more generally clarifies relational realism's explanatory model; indeed everything that Kiser and Hechter have said against the former would apply equally to the latter (Gould 1981, 1988, 1989; Hull 1988; Kauffman 1993). Evolutionary biology is narrative, historical, and lacks true invariant laws, yet is now understood to be the foundation of biology (along with genetic theory). It is, however, full of causal mechanisms and generalizing theory that explain how things happen and happened (Kauffman 1993). There are no laws for the extinction of species, for example, but there are indeed mechanisms: a comet hits and kills the dinosaurs thousands of years later. It was not, moreover, any intrinsic causal or essential property of the poor dinosaurs that made them extinct, nor anything intrinsic to comets that they extinguish dinosaurs; it was just the bad luck of the dinosaurs to have been at the wrong place at the wrong time. Yet the absence of law or exogenous properties does not prevent us from recovering causal generalization by comparing historical sequences and episodes—or what Gould (1989) calls "replaying the tape." Gould still requires that we test for causal attribution and employs a theory of comparative causality that he calls "consilience"—in which many independent sources "conspire" to demon-

strate the causal impact of historical patterns (see also Ragin 1987). The comparative test is crucial: In explaining outcome Y by a hypothesized historical pattern X, consilience requires that if X's earlier stages had not occurred or had developed differently, then Y either would not have happened at all or would have happened so differently that it would require a different explanation (see also Miller 1987 on counterfactuals). The method shows that Y makes sense and can be explained rigorously as the outcome of the causal processes of A through C plus X; but "no law of nature enjoined" Y, and any variant of Y emerging from differently configured antecedents would "have been equally explicable, though massively different in form and effect" (Gould 1989, pp. 278, 282-83, 288). The similarities between Gould's consilience and the thickly temporal demands of path dependence and causal narrativity are instructive.

How Do We Know if a Theory Is True? Problem-Driven Pragmatist Reasoning

Kuhn challenged the accepted premise that knowledge is neither primarily theory driven (deductive) nor data driven (inductive) but problem driven, thus moving us beyond the limited dichotomy of induction versus deduction. His challenge calls into question the legitimacy of rational choice theorists' common accusation that the only alternative to theory-centric deductivism is what they call the "empiricism" of "naive inductivism." Recent developments in science studies have taken up Kuhn's challenge and so cast further doubts on such an unbalanced overly theory-centric view of science and the production of knowledge. Increasingly, we are learning that Kuhn was right when he suggested that philosophy's longstanding privileging of theory over history and practice was little more than an "image" (Kuhn 1970, p. 1) based on idealized, rather than historical reconstructions of science. What has emerged is strikingly different: Hacking (1983, 1992) emphasizes "intervening" over "representing," Pickering (1989, 1990, 1992a, 1993) stresses "the temporality of practice," Humphreys (1988) characterizes the best kind of inquiry as "empirical realism," and Bohman (1991) sums up the long overdue recognition of pragmatism as "the new logic of social science." All follow Kuhn in rejecting philosophical images in favor of reconstructing the actual histories and practices of inquiry and knowledge production in science and the social sciences. The shift entails a move away from rationalist and a priori criteria for confirmation to "problems as they emerge in the theories and research of scientific practice itself" (Bohman 1991, p. 53). A more experimental and practice-centered view of theory confirmation than allowed by the either/or of deduction versus induction has thus gained scholarly recognition. Along with the emphasis being placed on the practical ele-

ments of inquiry has also come an overdue recognition of the diverse criteria that can justify (and have justified) theories (see esp. Pickering 1992). We now know that testing theories by observing only those phenomena designated by a preexisting hypothesis is only one among a wide variety of avenues along which science can lead us toward path-breaking knowledge. The relative import of general theory-driven logic has indeed been radically exaggerated (Galison 1987, 1989; Gooding et al. 1989; Gooding 1992; Pickering 1989, 1990, 1992a, 1993).

Kuhn prefigured the new trend. Beneath the false alternatives of induction and deduction he pointed to evidence of the more common practice of "articulation." Acknowledging the significance of this alternative would, I believe, advance our understanding of scientific development. Hacking (1983) has been most influential in elaborating the notion of articulation. He adopts Kuhn's term to explain how much more complex is the bridge between hypotheses and data than the limited either/or of deduction versus induction (Hacking 1983, pp. 213–16). Articulation more accurately represents the practice by which scientists "create the phenomena which then become the centerpieces of theory" (p. 220). Against a disproportionate theory-centrism, Hacking translates Kuhn's term to mean the history "not of what we think but of what we do . . . because . . . reality has more to do with what we do in the world than with what we think about it" (p. 17). Following Hacking, the historically forged notion of articulation has been widely developed into a full-blown alternative practice-centered conception of theory construction in science (see esp. Pickering 1992b).

In social science, Stinchcombe (1978) foreshadowed this trend when he declared that "the dilemma between synthetic reasoned generality, tested against the facts, and historical uniqueness, a portrait of the facts, is a *false dilemma*. The way out of the dilemma is that portraits of the facts, combined with an intellectual operation of carefully drawn analogies, are roads to generality" (pp. 115–16; emphasis added). Western (1997, pp. 10–12) has summarized cogently the problem-driven approach as one that involves (a) a willingness to make strong use of *substantive* knowledge to guide assumptions, (b) admission of the heterogeneous sources of uncertainty associated with all data, and (c) building models to fit our substantive problems and *not the other way around*. The problem-driven and practice-centered approach has a new credibility not reducible to either induction or deduction; a disproportionate theory-driven adherence to deductivism is no longer feasible (see also Humphreys 1988).

Theorizing convincingly about mechanisms, then, is a task requiring neither pure induction nor pure deduction but one that demands devising diverse and creative ways (e.g., abduction or articulation) to answer the question of whether the theoretical entity being hypothesized can actually be demonstrated to have a relational effect on a specific problem: Are, for

example, the mechanisms triggered by Putnam's (1993) differing degrees of social capital really viable theorizations for explaining varying patterns of democratization? Only through research practices deploying substance-driven comparisons and problem-driven manipulations can we get an answer to whether a "theoretical entity" such as social capital actually bears what Abbott (1996, p. 873) calls "causal authority" when he asks: Does something have an "independent standing as a site of causation, as a thing with consequences . . . [an ability] to create an effect on the rest of the social process that goes beyond effects . . . merely transmitted through the causing entity from elsewhere?"

TAKING STOCK

I began by posing a puzzling paradox: How is it possible that Kiser and Hechter's attack on historical sociology could have its epistemological roots in a recognizable post-Kuhnian philosophy of science and yet be at militant odds with the very incorporation of history that Kuhn so eloquently advocated? Let me briefly crystallize the three-part argument through which I have addressed this paradox. First, I suggested that there have been radically competing interpretations of Kuhn's argument and, from these, it is possible to draw two post-Kuhnian branches of realism—one theoretical, the other relational and pragmatic. Of these two routes out of Kuhn, the long dominant one has been a theory-centric antipositivism that converged with Popper's deductivism—even while it adamantly refused standard deductivism's ultimate foundations in the language of observation. Deeply opposed to positivism's antirealism, the theoretical realist route out of Kuhn gave a new lease on life to an age-old rationalist epistemology. Given Kuhn's fundamental demonstration of the centrality of history for epistemology, it is a great irony that among the major casualties of theoretical realism is history—once again (qua Kiser and Hechter 1991) relegated to the status of "particularistic," "antitheoretical" epistemological Other to "real" theory.

The paradox that I have pointed to is that although seemingly anticipated by Friedman's (1953) instrumentalist defense of unreal "as if" assumptions, rational choice avows a *non*instrumental belief in the *reality* of its assumptions, and this is a feat only made possible by the theory-centric version of the Kuhnian revolution. Once the "theory-laden" characterization of observation had gained hegemony, not only induction but even more, "mere history" became tainted with the empiricist label. With this move the way was wide open for theoretical realism to claim a higher form of epistemological privilege for the power of abstract reason, logic, and a priori principles. Worries about empirical or historical grounding are met with dismissals of "empiricism" and "antitheoreticism," such as

those that drive Kiser and Hechter's argument. Rational choice, the theoretical position from which Kiser and Hechter launch their attack, is the current foremost standard bearer of theoretical realism's return to the dismissive attitude toward history that precipitated Kuhn's original challenge. Stating, for example, that narrative is only useful for purely "illustrative" purposes, Kiser and Hechter invoke memories of Kuhn's critique of an epistemology that limits the value of history to "chronology" and "anecdote." Hence the paradox comes full circle in their assault on historical sociology.

Second, I have proposed a reading of Kuhn that allows us to develop an alternative branch of realism, one that follows his challenge to liberate history from its demeaned status as descriptor, anecdote, and mere illustration. Advocating a historical epistemology, I have tried to show that theory must be considered historical in a double sense: (1) What makes a theory true does not and cannot exist outside of the spectrum of historically conceivable questions of its time; it is only the questions we ask that bring into existence those theories that compete to be answers and it is only questions that transform random facts into confirming evidence. And (2) for any theory to provide an explanation requires a structure of causal linkages (mechanisms) and what Tilly (1994) calls a thick sense of history. Concomitantly, this means that history could no longer be caricatured as random particularities and excluded from the domain of true theory. Instead he showed history to bear within it the accumulation of questions that drive the production of knowledge; in turn, history gains epistemological stature as the stuff of sequences, processes, and relations that we can arrange into causal interventions and explanations based on path dependency and causal narrativity.

The third dimension of my argument has been to make explicit and to challenge the epistemological foundations of Kiser and Hechter's rational choice theory. With respect to their criteria for what counts as an "adequate explanation," Kiser and Hechter: (a) use a tautological methodology in which they impute causal mechanisms to the case at hand from an ontology already comprised of causal mechanisms (self-propelling agential intentionality); (b) they confound the covering law model's rules for theory justification with rational choice's realist reliance on inferred theoretical entities—hence mixing realism and antirealism with respect to observation rules; and (c) they create a hybrid standard for explanatory adequacy by conflating the abstract conjunctions of predictive laws with true causality—thus preempting the very process centeredness of the causal mechanisms that they advocate. In considering their criteria for justification, moreover, we deserve better reasons for why we should be convinced than mere appeals to the elegance and parsimony of their preferred general theory of rational choice. And with respect to their ontology, the problem

is not the individualism, the attributed rationality, or the micro level of their basic unit of analysis, but their embrace of a causal ontic methodology: in advance of inquiry they ascribe to the individual agent exogenous a priori self-propelling causal properties that necessarily (deterministically) and universally will cause action. There is nothing wrong with an explanation based on intentionality; its causal powers in any given theory however, must be the result of research and inquiry, not of causal ontological postulates. A social science that presumes precisely that which should have to bear the test of demonstration—namely, the causal factors that best explain a given outcome or problem—is a social science that eliminates the contingency now recognized to be essential in causal theories (Lieberson 1991).

Kiser and Hechter, in sum, thus present a mixture of pre-Kuhnian deductivism and covering law model logic with a post-Kuhnian theory-centric causal ontology and commitment to causal mechanisms (see, e.g., 1991, p. 6). In the first instance, they rely explicitly on the authority of Hume. The price of starting with Hume, however, is that Hume insisted that experience gives us only observable correlations ("constant conjunctions"), and so little can be said about causal laws (except in terms of statistical probability) and nothing at all about causal mechanisms (except that none are observable). Since Kiser and Hechter are not willing to pay this price, they turn to theoretical realism and its claim that the unobservable can be inferred from general theory. This gives them license to build explanations based entirely on the premise of a universal causal ontology and a priori axioms of agents-as-mechanisms—explanations that are confirmed when their implications conform to the standards of justification set by theoretical realism.

CONCLUSION

Kuhn knew the long-standing caricature of the epistemological dichotomy between history and theory would not die easily; he even suggested that the "main obstacle" to the acceptance of his historical epistemology would likely be similar to the attacks on historically based theory that confronted Darwin's *Origins of Species* (Kuhn 1970, pp. 171–72). That Kiser and Hechter's chief accusation and interpretation of historical sociology is a "relativist" one shows just how prescient Kuhn was. Almost 20 years after *The Structure of Scientific Revolutions*, Stinchcombe (1978) recommended that the social sciences follow a path similar to Kuhn's: "There is nothing so likely to cause a misreading of this argument [history vs. theory is a false dichotomy] as twisting it back into the psychology of generality that it rejects. The Kantian versus Nietzsche (or positivist vs.

Dilthey) version of what epistemology is all about is so embedded in the historical origins of social science that an argument based on the supposition that this is the wrong question has a hard time trying to say what it is about" (p. 116). The recent "historic turn" in science studies and the philosophy of science would seemingly make it increasingly difficult to play an image-based and idealized version of the science card in criticizing history. Yet Kiser and Hechter's attack on historical sociology as "antitheoretical" and "relativist" is but the latest manifestation of Stinchcombe's observation of just how hard it is to say what an historical epistemology "is about." Trundling out these same tired caricatures of both history and theory, they prove just how intractable the false dichotomy is, but one as indefensible today now as it has ever been.

Paradoxically, then, two leading historical sociologists are taking us right back to the antihistorical position that Kuhn originally challenged—free now, apparently, as post-Kuhnians from the discipline of observation with which positivism was burdened. But positivism burdened itself for good reason with the discipline of the empirical—to escape the potential of epistemological tyranny, arbitrary dogma, and prejudice. In the absence of this discipline, the appeal to "theory" can be as much a weapon of mystification as the blunt sword of "empiricism." The theoretical world that rational choice theory inhabits may well be, as Kiser and Hechter claim, "parsimonious," "bold," even elegant; but we need, nonetheless, justification for believing in it.

Kiser and Hechter are quick to scrutinize and to condemn the errors of historical sociologists. They have been less quick to examine and to question how their own epistemological strictures have histories, histories not of refinement towards greater absolute truth, but rather histories of deep contestation—histories not unlike those that both they and we study. Unfortunately, the histories that constitute some of our most important and significant explanatory rules are often completely invisible (Lieberson 1991); they have been naturalized to the point where Kiser and Hechter are able to confound general standards for what counts as theory and knowledge with the specificities of rational choice theory—and thus privilege theoretical realism over "mere history." They have in effect preempted the benefits of deliberation among rival positions (clarifying pre-suppositions, broadening purview, etc.) and so averted the difficult questions inherent in the multiplicity of intents, purposes, and radically divergent assumptions characteristic of the social sciences today. But the consequences of their imperial claims are ironic: Kiser and Hechter have forcefully insisted on the postpositivist truisms that all observations are theory laden and that historical accounts are organized by theoretical categories. But perhaps Kiser and Hechter should ponder whether their own

evaluative criteria are themselves not history laden, and thus contingent and embedded in a larger field of less angelic, more pluralist, epistemological possibilities.

REFERENCES

- Abbott, Andrew 1988. "Transcending General Linear Reality." *Sociological Theory* 6:169-86.
- . 1992a. "What Do Cases Do? Some Notes on Activity in Sociological Analysis." Pp 53-82 in *What Is a Case?* edited by Charles Ragin and Howard Becker New York: Cambridge University Press
- . 1992b. "From Causes to Events Notes on Narrative Positivism." *Sociological Methods and Research* 20 (4) 428-55.
- . 1993. "Measure for Measure." *Journal of Mathematical Sociology* 18 203-14.
- . 1995. "Sequence Analysis: New Methods for Old Ideas." *Annual Review of Sociology* 21:93-113.
- . 1996. "Things of Boundaries." *Social Research* 62:857-82.
- Adams, Julia. In press. "Culture in Rational Choice Theories of State Formation." In *State/Culture: State Formation after the Cultural Turn*, edited by George Steinmetz. Ithaca, N.Y.: Cornell University Press.
- Alexander, Jeffrey C. 1982. *Positivism, Presuppositions, and Current Controversies*. Vol. 1 of *Theoretical Logic in Sociology*. Berkeley and Los Angeles: University of California Press.
- Almond, Gabriel A. 1988. "Separate Tables: Schools and Sects in Political Science." *PS: Political Science and Politics* 21:828-42
- Aminzade, Ron. 1992. "Historical Sociology and Time." *Sociological Methods and Research* 20 (4): 456-80.
- Anderson, Elizabeth. 1993. *Values in Economics* Cambridge, Mass.: Harvard University Press.
- Apel, Karl-Otto. 1994. *Understanding and Explanation*. Cambridge, Mass.: MIT Press.
- Arthur, Brian W. 1988. "Self-Reinforcing Mechanisms in Economics." In *The Economy as an Evolving Complex System*, edited by P. Anderson, K. Arrow, and D. Pines Reading, Mass.: Addison Wesley.
- . 1989. "Competing Technologies, Increasing Returns, and Lock-in by Historical Events." *Economic Journal* 99:116-31.
- Bates, Robert, et al. 1998. Introduction to *Analytic Narratives*, edited by Robert Bates, Avner Greif, Margaret Levi, Jean-Laurent Rosenthal, and Barry Weingast Cambridge: Cambridge University Press.
- Bearman, Peter. 1993. *Relations into Rhetorics*. New Brunswick, N.J.: Rutgers University Press
- Bell, Daniel. 1980. "Models and Reality in Economic Discourse." Pp. 46-80 in *The Crisis in Economic Theory*, edited by Daniel Bell and Irving Kristol. New York: Basic Books
- Bhaskar, Roy 1979. *A Realist Theory of Science*, 2d ed. Brighton: Harvester.
- . 1986. *Scientific Realism and Human Emancipation*. London: Verso.
- . 1989. *The Possibility of Naturalism*, 2d ed Hertfordshire: Harvester.
- Block, Fred. 1991. *Postindustrial Possibilities: A Critique of Economic Discourse*. Berkeley and Los Angeles: University of California Press.
- Block, Fred, and Margaret Somers. 1984. "Beyond the Economistic Fallacy: The Holistic Social Science of Karl Polanyi." Pp 47-84 in *Vision and Method in Histor-*

- ical Sociology*, edited by Theda Skocpol. New York: Cambridge University Press.
- Bohman, James. 1991. *The New Logic of Social Science*. Cambridge: Polity Press.
- Boyd, Richard N. 1973. "Realism, Underdetermination, and a Causal Theory of Evidence." *Nous* 7:1-12.
- Brelger, Ronald L. 1991. *Explorations in Structural Analysis: Dual and Multiple Networks of Social Interaction*. New York: Garland.
- Burt, Ronald S. 1992. *Structural Holes: The Social Structure of Competition*. Cambridge, Mass.: Harvard University Press.
- Calhoun, Craig. 1995. *Critical Social Theory*. Oxford: Blackwell.
- Camic, Charles, and Yu Xie. 1994. "The Statistical Turn in American Social Science: Columbia University, 1890-1915." *American Sociological Review* 59:773-805.
- Cartwright, Nancy. 1983. *How the Laws of Physics Lie*. Oxford: Clarendon Press.
- . 1989. *Nature's Capacities and their Measurement*. Oxford: Clarendon.
- Cassirer, Ernst. 1953. *Substance and Function*, translated by W. Curtis Swabey. New York: Dover.
- Coleman, James S. 1986. "Social Theory, Social Research, and a Theory of Action." *American Journal of Sociology* 91 (6) 1309-35.
- . 1990. *Foundations of Social Theory*. Cambridge, Mass.: Harvard University Press.
- Collier, Andrew. 1994. *Critical Realism*. New York: Verso.
- Collingwood, R. G. (1939) 1970. *An Autobiography*. Oxford: Oxford University Press.
- Collins, Randall. 1981. "On the Microfoundations of Macrosociology." *American Journal of Sociology* 86:984-1014.
- David, Paul. 1985. "Clio and the Economics of QWERTY." *American Economic Review* 75:332-37.
- Dessler, David. 1989. "What's at Stake in the Agent-Structure Debate?" *International Organisation* 43:441-73.
- Dray, W. H. 1957. *Laws and Explanations in History*. London: Oxford University Press.
- Duhem, Pierre. (1906) 1954. *The Aim and Structure of Physical Theory*. Princeton, N.J.: Princeton University Press.
- Elster, Jon. 1983. *Explaining Technical Change*. Cambridge: Cambridge University Press.
- . 1989. *Nuts and Bolts for the Social Sciences*. Cambridge: Cambridge University Press.
- Emirbayer, Mustafa. 1997. "Manifesto for a Relational Sociology." *American Journal of Sociology* 103 (2): 281-317.
- Entrikin, J. Nicholas. 1991. *The Betweenness of Place: Towards a Geography of Modernity*. Baltimore: Johns Hopkins University Press.
- Florina, Morris P., and Kenneth A. Shepsle. 1982. "Equilibrium, Disequilibrium, and the General Possibility of a Science of Politics." In *Political Equilibrium*, edited by Peter C. Ordeshook and Kenneth A. Shepsle. The Hague: Kluwer-Nijhoff.
- Friedman, Jeffrey, ed. 1996. *The Rational Choice Controversy*. New Haven, Conn.: Yale University Press.
- Friedman, Milton. 1953. "The Methodology of Positive Economics." In *Essays in Positive Economics*, edited by Milton Friedman. Chicago: University of Chicago Press.
- Fuller, Steve. 1992. "Being There with Thomas Kuhn: A Parable for Postmodern Times." *History and Theory* 31:241-75.
- Galison, Peter. 1987. *How Experiments End*. Chicago: University of Chicago Press.
- . 1989. "The Trading Zone: Coordinating Action and Belief." Manuscript. Stanford University.
- Giddens, Anthony. 1984. *The Constitution of Society*. Cambridge: Polity Press.

- Gooding, David 1992. "Putting Agency Back into Experiment." Pp. 65–112 in *Science as Practice and Culture*, edited by A. Pickering. Chicago: University of Chicago Press.
- Gooding, David, T. J. Pinch, and S. Schaffer, eds. 1989. *The Uses of Experiment*. Cambridge: Cambridge University Press.
- Gould, Stephen Jay. 1981. *The Mismeasure of Man*. New York: Norton.
- . 1988. "Mighty Manchester." *New York Review of Books*, October 27.
- . 1989. *Wonderful Life*. Cambridge, Mass.: Harvard University Press.
- Green, Donald, and Ian Shapiro. 1994. *The Pathologies of Rational Choice Theory*. New Haven, Conn.: Yale University Press.
- Griffin, Larry 1992 "Temporality, Events, and Explanation in Historical Sociology" *Sociological Methods and Research* 20:403–27.
- . 1993 "Narrative, Event-Structure Analysis, and Causal Interpretation in Historical Sociology." *American Journal of Sociology* 98:1094–1133.
- Gutting, Gary, ed. 1980. *Paradigms and Revolutions*. Notre Dame, Ind.: University of Notre Dame Press.
- Hacking, Ian. 1983 *Representing and Intervening*. Cambridge: Cambridge University Press.
- . 1990a. *The Taming of Chance*. Cambridge: Cambridge University Press.
- . 1990b. "Two Kinds of 'New Historicism' for Philosophers." *New Literary History* 21:343–64.
- . 1992. "The Self-Vindication of the Laboratory Sciences." Pp. 29–64 in *Science as Practice and Culture*, edited by A. Pickering. Chicago: University of Chicago Press.
- Hechter, Michael 1992. "Rational Choice Theory and Historical Sociology." *International Social Science Journal* 133:367–73.
- Hempel, Carl G. (1942) 1959. "The Function of General Laws in History." Pp. 344–56 in *Theories of History*, edited by P. Gardiner. New York: Free Press.
- . 1962 "Rational Action." *Proceedings of the American Philosophical Association*.
- . 1965. "Aspects of Scientific Explanation." Pp. xx in *Aspects of Scientific Explanation*, edited by Carl G. Hempel. New York: Free Press.
- Hesse, Mary. 1980. *Revolutions and Reconstructions in the Philosophy of Science*. Bloomington: Indiana University Press.
- Hirschman, Albert. 1984. "Against Parsimony." *American Economic Papers and Proceedings*, 89–96.
- Hollinger, David. 1980 "T. S. Kuhn's Theory of Science and Its Implications for History." Pp. 195–222 in *Paradigms and Revolutions*, edited by Gary Gutting.
- Hollis, Martin. 1994. *The Philosophy of Social Science*. Cambridge: Cambridge University Press.
- Hollis, Martin, and Steven Lukes, eds. 1982. *Rationality and Relativism*. Cambridge, Mass.: MIT Press.
- Hollis, Martin, and Steve Smith. 1990. *Explaining and Understanding International Relations*. Oxford: Oxford University Press.
- . 1991. "Beware of Gurus: Structure and Action in International Relations." *Review of International Studies* 17:393–410.
- . 1994 "Two Stories about Structure and Agency." *Review of International Studies* 20:241–51.
- Holmes, Oliver Wendell (1881) 1963. *The Common Law*, edited by Mark DeWolfe. Cambridge, Mass.: Harvard University Press.
- Hull, David. 1977 "The Ontological Status of Species." P. 99 in *Foundational Problems in the Special Sciences*. Pt. 2 of *Proceedings of the Fifth International Congress of Logic, Methodology and the Philosophy of Science*. London, Ontario, Canada-1975, edited by R. Butts and J. Hintikka. Dordrecht: Reidel.

- . 1978. "A Matter of Individuality." *Philosophy of Science* 45:335–360.
- . 1981. "Discussion Kitts and Kitts and Caplan on Species." *Philosophy of Science* 48:147.
- . 1988. *Science as a Process: An Evolutionary Account of the Social and Conceptual Development of Science*. Chicago: University of Chicago Press.
- Hume, David. 1938 *An Abstract of a Treatise of Human Nature*, edited by J. M. Keynes and P. Straffa. Cambridge: Cambridge University Press.
- Humphreys, Paul. 1988. "Causal, Experimental, and Structural Realisms." *Midwest Studies in Philosophy* 12:241–52.
- . 1989. *The Chances of Explanation*. Princeton, N.J.: Princeton University Press.
- Kauffman, Stuart A. 1993. *The Origins of Order: Self-Organization and Selection in Evolution*. New York: Oxford University Press.
- King, Gary, Robert O. Keohane, and Sidney Verba. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton, N.J.: Princeton University Press.
- Kiser, Edgar. 1996. "Narratives and Rational Choice." *Politics and Society*
- Kiser, Edgar, and Michael Hechter. 1991. "The Role of General Theory in Comparative-Historical Sociology." *American Journal of Sociology* 97:1–30.
- Koyre, Alexandre. 1957. *From the Closed World to the Infinite Universe*. Baltimore: Johns Hopkins University Press.
- Kuhn, Thomas. (1962) 1970. *The Structure of Scientific Revolutions*, 2d ed. Chicago: University of Chicago Press.
- Lakatos, Imre. 1970. "Falsification and the Methodology of Scientific Research Programmes." In *Criticism and the Growth of Knowledge*, edited by I. Lakatos and A. Musgrave. Cambridge: Cambridge University Press.
- . 1976. *Proofs and Refutations: The Logic of Mathematical Discovery*. Cambridge: Cambridge University Press.
- . 1978. *The Methodology of Scientific Research Programs*. Cambridge: Cambridge University Press.
- Laumann, Edward O., Peter V. Marsden, and David Prensky. 1983. "The Boundary Specification Problem in Network Analysis." Pp. 18–34 in *Applied Network Analysis: A Methodological Introduction*, edited by Ronald S. Burt and Michael J. Minor. Beverly Hills, Calif.: Sage.
- Lieberson, Stanley. 1991. "Einstein, Renoir and Greeley: Some Thoughts about Evidence in Sociology." *American Sociological Review* 57:1–15.
- Lloyd, Christopher. 1986. *Explanation in Social History*. Oxford: Basil Blackwell.
- . 1993. *Structures of History*. Oxford: Basil Blackwell.
- Longino, Helen. 1990. *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry*. Princeton, N.J.: Princeton University Press.
- MacIntyre, Alasdair. 1980. "Epistemological Crises, Dramatic Narrative, and the Philosophy of Science." Pp. 54–74 in *Paradigms and Revolutions*, edited by G. Gutting. Notre Dame, Ind.: University of Notre Dame Press.
- Mann, Michael. 1986. *Sources of Social Power*, vol. 1. New York: Cambridge University Press.
- McDonald, Terrence. 1996. "The Historic Turn in the Human Sciences." Ann Arbor: University of Michigan Press.
- Miller, Richard. 1987. *Fact and Method*. Princeton, N.J.: Princeton University Press.
- Newton-Smith, W. (1958) 1978. "The Underdetermination of Theory by Data." *Proceedings of the Aristotelian Society*, suppl., vol. 52.
- Outhwaite, William. 1987. *New Philosophies of Social Science*. New York: St. Martin's Press.
- Padgett, John F., and Christopher K. Ansell. 1993. "Robust Action and the Rise of the Medici, 1400–1434." *American Journal of Sociology* 98:1259–1319.
- Pickering, Andrew. 1989. "Living in the Material World: On Realism and Experimentation."

- tal Practice." Pp. 275-97 in *The Uses of Experiment*, edited by David Gooding, T J Pinch, and S. Schaffer. Cambridge: Cambridge University Press.
- . 1992a. "From Science as Knowledge to Science as Practice." Pp. 1-28 in *Science as Practice and Culture*, edited by A. Pickering. Chicago: University of Chicago Press.
- , ed. 1992b. *Science as Practice and Culture*. Chicago: University of Chicago Press.
- . 1993. "The Mangle of Practice. Agency and Emergence in the Sociology of Science." *American Journal of Sociology* 99 (3): 559-89.
- Pizzorno, Alessandro. 1986. "Some Other Kinds of Otherness." Pp. 355-73 in *Development, Democracy, and the Art of Trespassing: Essays in Honor of Albert O. Hirschman*, edited by Alejandro Foxley, Michael S. McPherson, and Guillermo O'Donnell. Notre Dame, Ind.: University of Notre Dame Press.
- . 1991. "On the Individualistic Theory of Social Order." Pp. 209-31 in *Social Theory for a Changing Society*, edited by Pierre Bourdieu and James S. Coleman. Boulder, Colo.: Westview Press.
- . 1995. "On the Microfoundations of Social Change." Manuscript. European University Institute, Florence.
- Polyanyi, Michael. 1957. *Personal Knowledge*. Chicago: University of Chicago Press.
- Popper, Karl (1934) 1959. *The Logic of Scientific Discovery*. New York: Basic Books.
- . 1950. *The Open Society and Its Enemies*. Princeton, N.J.: Princeton University Press.
- Putnam, Hilary. 1975. *Mind, Language and Reality*. Cambridge: Cambridge University Press.
- . 1997. "A Half Century of Philosophy, Viewed from Within." *Daedalus* 12 (1): 175-208.
- Putnam, Robert. 1993. *Making Democracy Work*. Princeton, N.J.: Princeton University Press.
- Quadagno, Jill, and Stan J. Knapp. 1992. "Have Historical Sociologists Forsaken Theory? Thoughts on the History/Theory Relationship." *Sociological Methods and Research* 20 (4): 481-507.
- Quine, Willard Van Orman. 1963. "Two Dogmas of Empiricism." In *From a Logical Point of View*. New York: Harper & Row.
- . 1969. *Ontological Relativity and Other Essays*. New York: Columbia University Press.
- Ragin, Charles. 1987. *The Comparative Method*. Berkeley and Los Angeles: University of California Press.
- Ragin, Charles, and Howard Becker. 1992. *What Is a Case? Exploring the Foundations of Social Inquiry*. Cambridge: Cambridge University Press.
- Reichenbach, Hans. 1951. *The Rise of Scientific Philosophy*. Berkeley and Los Angeles: University of California Press.
- Rorty, Richard. 1979. *Philosophy and the Mirror of Nature*. Princeton, N.J.: Princeton University Press.
- Rosenberg, Alexander. 1988. *Philosophy of Social Science*. Oxford: Clarendon.
- Salmon, Wesley. 1984. *Scientific Explanation and the Causal Structure of the World*. Princeton, N.J.: Princeton University Press.
- Scott, John. 1991. *Social Network Analysis: A Handbook*. Newbury Park, Calif.: Sage.
- Sewell, William H., Jr. 1996. "Three Temporalities: Toward a Sociology of the Event." In *The Historic Turn in the Human Sciences*, edited by Terrence J. McDonald. Ann Arbor: University of Michigan Press.
- Shepsle, Kenneth A. 1990. "The Nature of Contemporary Political Science." *PS: Political Science and Politics* 23:40-42.

- Skocpol, Theda. 1979. *States and Social Revolutions*. Cambridge: Cambridge University Press.
- . 1984. "Emerging Agendas and Recurrent Strategies in Historical Sociology." Pp. 356–91 in *Vision and Method in Historical Sociology*, edited by Theda Skocpol. New York: Cambridge University Press.
- . 1994. "Reflections on Recent Scholarship about Social Revolutions and How to Study Them." Pp. 301–43 in *Social Revolutions in the Modern World*, edited by Theda Skocpol. Cambridge: Cambridge University Press.
- Skocpol, Theda, and Margaret R. Somers. 1980 "The Uses of Comparative History in Macrosocial Inquiry." *Comparative Studies in Society and History* 22:174–97.
- Smelser, Neil. 1992. "The Rational Choice Perspective." *Rationality and Society*, 4: 381–410.
- Somers, Margaret R. 1993 "Citizenship and the Place of the Public Sphere: Law, Community and Political Culture in the Transition to Democracy." *American Sociological Review* 58:587–620.
- . 1994a. "Narrative and the Constitution of Identity: A Relational and Network Approach." *Theory and Society* 23 (5): 605–50.
- . 1994b. "Rights, Relationality, and Membership: Rethinking the Making and Meaning of Citizenship." *Law and Social Inquiry* 19 (1): 63–112
- . 1996a. "Class Formation and Capitalism." *European Journal of Sociology* 37 (1): 180–202.
- . 1996b. "Where Is Sociology after the Historic Turn? Knowledge Cultures and Historical Epistemologies." In *The Historic Turn in the Human Sciences*, edited by Terrence J. McDonald. Ann Arbor: University of Michigan Press.
- Somers, Margaret, and Gloria Gibson. 1994. "Reclaiming the Epistemological 'Other': Narrative and the Social Constitution of Identity." Pp. 37–99 in *Social Theory and the Politics of Identity*, edited by Craig Calhoun. Oxford: Basil Blackwell.
- Stark, David. 1992. "Path Dependence and Privatization Strategies in East Central Europe." *East European Politics and Societies* 6:17–54.
- Steinmetz, George. 1992. "Reflections on the Role of Social Narratives in Working-Class Formation." *Social Science History* 16 (3): 489–516.
- . 1993. *Regulating the Social: The Welfare State and Local Politics in Imperial Germany*. Princeton, N.J.: Princeton University Press.
- . 1998. "Critical Realism and Historical Sociology." *Comparative Studies in Society and History* 40 (1): 170–86.
- Steinmo, Sven, Kathleen Thelen, and Frank Longstrech. 1992. *Structuring Politics*. Cambridge: Cambridge University Press.
- Stinchcombe, Arthur L. 1978. *Theoretical Methods in Social History*. New York: Academic Press.
- . 1991. "The Conditions of Fruitfulness of Theorizing about Mechanisms in Social Science." Pp. 24–41 in *Social Theory and Social Policy: Essays in Honor of James S. Coleman*, edited by Aage B. Sorensen and S. Spillerman. Westport, Conn.: Praeger.
- . 1992. "Simmel Systematized: James S. Coleman and the Social Forms of Purposive Action in His Foundations of Social Theory." *Theory and Society* 21:183–202.
- . 1993. "Freedom and Oppression of Slaves in the Eighteenth-Century Caribbean." *American Sociological Review* 59:911–29.
- . 1995. *Sugar Island Slavery in the Age of Enlightenment: The Political Economy of the Caribbean World*. Princeton, N.J.: Princeton University Press.
- Tilly, Charles. 1984. *Big Structures, Large Processes, Huge Comparisons*. New York: Russell Sage.
- . 1988. "Misreading, then Rereading, Nineteenth-Century Social Change." Pp

- 332–358 in *Social Structures: A Network Approach*, edited by B. Wellman and S. D. Berkowitz. Cambridge: Cambridge University Press
- . 1990. *Coercion, Capital, and European States, A.D. 990–1990*. Cambridge, Mass.: Basil Blackwell.
- . 1994. "The Time of States" *Social Research* 61:269–95.
- . 1995a. "Macrosociology, Past and Future." Pp. 1–3 in *The Relational Turn in Macrosociology: A Symposium*. Working Paper no. 215. New School for Social Research, Center for Studies of Social Change
- . 1995b. *Popular Contention in Great Britain: 1754–1828*. Cambridge, Mass.: Harvard University Press.
- . 1995c. "To Explain Political Processes." *American Journal of Sociology* 100 (6): 1594–1610
- Turner, Jonathan. 1992. "The Promise of Positivism." Pp. 156–78 in *Postmodernism and Social Theory*, edited by Steven Seidman and David G. Wagner. Cambridge, Mass.: Blackwell
- Wacquant, Loïc J. D. 1993. "Positivism." Pp. 495–98 in *Twentieth-Century Social Thought*, edited by William Outhwaite and Tom Bottomore. Oxford: Blackwell.
- Wallerstein, Immanuel. 1991. *Unthinking Social Science*. Cambridge, Mass.: Blackwell.
- Wellman, Barry. 1988. "Structural Analysis: From Method and Metaphor to Theory and Substance." Pp. 19–61 in *Social Structures: A Network Approach*, edited by Barry Wellman and S. D. Berkowitz. Cambridge: Cambridge University Press
- Western, Bruce. 1997. *Between Class and Market: Postwar Unionization in the Capitalist Democracies*. Princeton, N.J.: Princeton University Press.
- White, Harrison C. 1992. *Identity and Control: A Structural Theory of Social Action*. Princeton, N.J.: Princeton University Press.
- . 1993. "Strategizing and Mobilization." Manuscript. Columbia University
- . 1996. "Network Switchings and Bayesian Forks: Reconstructing the Social and Behavior Sciences." *Social Research* 62:1035–63.
- White, Harrison C., Scott A. Boorman, and Ronald L. Breiger. 1976. "Social Structure from Multiple Networks I: Blockmodels of Roles and Positions." *American Journal of Sociology* 81 (4): 730–80.
- Zald, Mayer. 1991. "Sociology as a Discipline: Quasi-Science and Quasi-Humanities." *American Sociologist* 22 (3,4): 165–87.

The Debate on Historical Sociology: Rational Choice Theory and Its Critics¹

Edgar Kiser
University of Washington

Michael Hechter
University of Arizona

In the past two decades, many sociologists have denied the usefulness of general theories in favor of more particularistic approaches to historical explanation, which makes it difficult to specify both the causal relations and the causal mechanisms that account for social outcomes. This article offers some philosophical and theoretical justifications for the use of general theory in historical analysis and contends that general theory guides the selection of facts, provides a source of generalizable causal mechanisms, facilitates the cumulation of knowledge across substantive domains, reveals anomalies that lead to new questions, and creates the conditions under which existing theories can be supplanted by superior ones. The authors further outline the concrete research practices that flow from their approach and discuss several empirical studies that exemplify these five advantages.

The recent revival of interest in historical sociology has renewed hoary old social scientific debates about methodology and the proper role of theory in historical explanation. Some time ago, in the pages of this journal, we discerned a trend toward a new style of research in comparative and historical sociology and pointed out some of its attendant liabilities (Kiser and Hechter 1991). Inductive methods were being endorsed over deductive ones, and general explanations of social outcomes were being denigrated in favor of those stressing complexity, uniqueness, and historical contingency. This tendency amounted to a considerable shift in emphasis:

¹ We thank Kathryn Baker, Terry Boswell, William Brustein, Sun-Ki Chai, Elisabeth Clemens, Debra Friedman, John R. Hall, Christine Horne, Ron Jepperson, Satoshi Kanazawa, Avery Kolers, and Misha Petrovic for helpful comments. We alone are responsible for the arguments presented herein. Direct correspondence to Edgar Kiser, Department of Sociology, University of Washington, 202 Savery Hall, Seattle, Washington 98195

whereas historical sociologists traditionally had insisted on distinguishing social science from history, now many sought to do away with the boundary between these disciplines. It seemed to us that the norms of history were coming to replace traditional sociological standards for assessing research.

We characterized this trend as troubling because it discourages adequate social explanations. Good explanations, we maintained, must specify both relations between causal factors (including models that indicate how the causal factors are related), and the mechanisms purporting to describe the process by which one causal factor influences the other. Since social mechanisms are unobservable, they cannot be derived from empirical observations; instead, they often are derived from general theories. Hence, we concluded, the increasingly fashionable attacks on general theory inexorably foster poor explanations of social outcomes. At the same time, we argued that some of the methodological practices in historical sociology were deficient.

To ground the argument more concretely, we briefly discussed state autonomy and policy formation—two favorite substantive topics in historical sociology. To this end, the article highlighted the research of Theda Skocpol and Michael Mann because it is widely acknowledged as exemplary.¹ We argued that one particular general theory—rational choice—improved their analyses by offering superior models of causal relations and by suggesting plausible causal mechanisms.²

Far from dismissing contemporary historical sociology *tout court*, our critique of historical sociology was selective.³ We challenged only method-

¹ In our 1991 article, we noted that "Skocpol and Mann have contributed a great deal to our understanding of state autonomy. They have demonstrated that states are potentially autonomous and that the extent of state autonomy varies across time and space. Both have identified structural factors related to variations in autonomy in particular historical settings, and each has made general statements about what sorts of factors are likely to influence state autonomy" (1991, p. 18). Our substantive aim in discussing state autonomy and policy formation was to use rational choice theory to "complete the project they have begun" (1991, p. 19). Further, we praised many aspects of Skocpol's work on revolutions: "Theda Skocpol's substantive studies of social revolutions and welfare states have combined important criticisms of existing theoretical perspectives with insightful empirical generalizations about important historical events. She has also been self-conscious and explicit about the methodological underpinnings of her work" (1991, p. 12).

² Somers (p. 723) claims that we view rational choice as the only current general theory. This is simply false. Our text (1991, pp. 19, 23, 16) referred to rational choice as "one source" of testable theoretical models and acknowledged "other general theories," including Marxism (see pp. 16, 19, 23). We still regard rational choice as the most promising general theory currently available and continue to focus on it throughout this discussion.

³ For example, our reservations about Stinchcombe refer only to his *Theoretical Methods in Social History* (1978). We approvingly cited his (1978) focus on causal mecha-

ologically questionable practices or discourses that denigrated the importance of general theory in social science. We never challenged the validity of the empirical generalizations advanced by historical sociologists and historians (contrary to Somers's claim above, p. 757).⁵ Our basic point was that if knowledge of general theory is one of sociologists' principal comparative advantages in conducting historical research, then this new counsel—that sociologists should consciously forswear this knowledge—is perverse.⁶

Since we were skeptical about the new wave in historical sociology, we expected the targets of our criticism to respond. There have been three responses to date.⁷ Quadagno and Knapp (1992) deny that historical sociologists reject theory; instead, they claim that historical sociologists have a very different notion of theory than we do. They advocate a "temporally grounded historical sociology," that stresses contingency and sequentiality, and focuses on "the theoretical value of narrative."⁸ Skocpol (1994) principally addresses our methodological criticisms and defends her own substantive analysis of revolutions. In what is by far the most detailed response, Somers (in this issue) elaborates on Quadagno and Knapp's epistemological criticisms and moves substantially beyond them. She also crit-

nisms (1991, p. 4), his (1968, pp. 19–20) discussion of testing theories (1991, p. 9), and his (1982) criticisms of Wallerstein (1991, p. 24). Although we were critical of some of Tilly's research, we also praised *The Vendee* (1964) and concurred with his (1975, pp. 11–12) argument that the best way to test causal statements is to follow cases over time (1991, p. 13). Our criticisms of Erikson (1966), Swanson (1960), and Wallerstein (1974) were confined to the complaint that they often failed to specify clear causal mechanisms.

⁵ Skocpol (1994, pp. 322–25) completely misconstrues our criticisms as referring to the substance of her historical arguments, rather than to her use of method and theory. Thus her complaint that we provide no superior alternative general theory of revolution—and cite only Taylor (1988) on peasant revolt—amounts to much ado about nothing. Taylor was cited only to make the point that Skocpol's (1979) explanation of revolution lacked the microfoundations necessary for providing a causal mechanism. Skocpol (1994, p. 324) herself has admitted that her argument was incomplete in precisely this respect.

⁶ We find exceedingly odd the notion that the social sciences should seek to emulate the humanities. Whatever one makes of the present state of the social sciences, many observers regard the intellectual health of the humanities as little short of perilous (Kernan 1997).

⁷ It thus seems strange that Somers would claim that we have "closed off debate" (p. 749).

⁸ Narrative analysis and rational choice are not necessarily at odds, however (Kiser 1996). The traditional mode of historical narrative has been intentional, relying on historical agents' presumed motives and beliefs. For McCloskey (1990), rational choice theory is one kind of narrative. The term often used by philosophers for intentional explanation, "folk psychology," reflects its origins in our intuitive ways of understanding others.

icizes our philosophical presuppositions and concludes with a broadside against rational choice theory.

A number of complex issues are broached in the debate. Here, we will address three main topics: philosophical presuppositions, rival views of theory, and the role of general theory—in particular, rational choice theory—in historical research.

Despite her claim that we share a commitment to epistemological realism, our philosophical differences with Somers are profound. Our own stance is in no way novel. We adhere to a standard version of realism as developed and elaborated by several philosophers of science (among them Kuhn [(1962) 1970], Popper [1994], Lakatos [1978], and Sayer [1992]; see also Searle 1995). We think that science progresses by testing clear general propositions and by attempting to resolve the anomalies that this testing process invariably reveals. Admittedly, this model of science sometimes rests on ambiguities concerning the criteria that make for a good test, or for the recognition of an anomaly. In the face of these ambiguities, Somers opts to throw the baby out with the bathwater: she advocates a novel and idiosyncratic set of philosophical foundations. Somers argues for a new historical epistemology and a new philosophy of science—she refers to it as relational realism—that derives from recent relativistic critiques of science. Her conception that science is nothing but a social construction shaped by particular historical circumstances offers little hope for the cumulation of knowledge. Accordingly, the first section of our response discusses these two views and refutes Somers's claim that our own philosophical foundations are incoherent.

Epistemological issues aside, we also have a very different conception of theory than do our critics, and we articulate those views in the second section of this discussion. On this score, as well, we make no great claims to originality. Our conception of theory closely resembles that found in advanced theory construction texts (Stinchcombe 1968; Cohen 1989). We subscribe to the basic Kuhnian notion that general theories provide a paradigm capable of producing a research program (see also Popper 1994). As such, general theories are composed of interrelated sets of assumptions, models, causal mechanisms, abstract scope conditions, and testable propositions. Even though this conception is extremely broad, our critics fault us for excessive narrowness. They espouse an alternative conception of theory that is more capacious by far. For them, theory merely consists of some loose orienting concepts or general categories. These are often constructed from historical evidence and untied to any overarching theoretical framework. If they exist at all, the causal mechanisms and scope conditions of their preferred theories are defined historically. The second section of our response therefore compares these rival conceptions of theory.

Which of these two rival conceptions is likely to produce superior research and lead to the cumulation of knowledge? To answer this question we must forsake the ethereal realm of epistemology for more earthly concerns about the practice of historical research. It would be profitable to compare our approach to Somers's at this terrestrial level, but she provides too few hints about the implications of her argument for carrying out historical research to do so.

The third section of this article shows how our preferred general theory, rational choice, can be used in historical analysis. To that end, it discusses the microfoundations and models used in rational choice, how arguments are tested, and how anomalies can be resolved. Several empirical examples serve to illustrate our approach, and we allude to some of the progress that has already been attained in this nascent research program. We conclude by summarizing the main points at issue in the debate and then offer several reasons for the superiority of our approach.

PHILOSOPHICAL PRESUPPOSITIONS

Our 1991 article focused on general theoretical and methodological issues in sociology, not on philosophy *per se*. Since the analysis was based on philosophical foundations, Somers is right to insist that we make these foundations more explicit, and that those foundations be coherent and consistent. We devoted little attention to these matters because we felt we had nothing original to offer—we presented no new epistemology, no new philosophy of science. Yet far from mixing realism and positivism, as Somers charges, we maintained a consistently realist position throughout.⁹ Our analysis is based on two epistemological assumptions: there is a social world existing outside the minds and sense perceptions of its participants, and mechanisms are responsible for generating causal relations. These are the main tenets of epistemological realism. Realism simply implies that the contours of this world are, at least in principle, objectively knowable. Because we often lack the capacity to perceive this real world fully, however, we must rely on concepts and theories that purport to describe its unobservable elements. This unfortunate, if necessary, reliance on unob-

⁹ Somers labels us "theoretical realists" to differentiate our position from the new form of realism she refers to as "relational realism." We follow standard practice in the philosophy of science and continue to use the term "realism," because her label of theoretical realism distorts our position. Despite our reluctance to delve into philosophy, we described our epistemological stance as realist (1991, p. 13; relying primarily on Lakatos [1978]) and made pains to distinguish it (1991, p. 6) from positivist views (e.g., Hempel 1959). Our treatment of causal explanation emphasized unobservable causal mechanisms and abstract models of causal relations, both of which are clearly realist concerns.

servables leads to a central difficulty with all realist explanations. If mechanisms are unobservable, then how can we be persuaded of their truth? For that matter, how can we adjudicate between rival explanations based on different mechanisms?

Mechanisms purport to describe the process by which a given cause brings about its effect. Although no hard and fast definition of them exists,¹⁰ mechanisms typically explain outcomes by invoking phenomena at a level of analysis lower than that of the outcome of concern (Stinchcombe 1991; Elster 1983; Bunge 1996). This means that one discipline's mechanisms are another's central research problems. For example, the mechanisms used in rational choice theory are intentional—actions are explained by the future states they are intended to bring about.¹¹ Whereas rational choice theorists (who are interested in accounting for macrolevel outcomes) invoke mechanisms based on purposive agents with transitive preferences, psychologists (who are interested in accounting for individual outcomes) invoke cognitions as mechanisms and conceive of the agent as an unwieldy macroscopic unit.

Philosophers of science are divided on how best to cope with causality. Positivists distrust causality and concentrate exclusively on observables (correlations or "constant conjunctions" between variables), but realists are willing to assume the existence of underlying causes that generate relations between observables. We argued that three criteria can help us assess explanations based on rival causal mechanisms: plausibility, reducing the time lag between cause and effect, and testing each mechanism's unique empirical implications (Kiser and Hechter 1991, pp. 6–8). The plausibility criterion—which resides not in individuals but in the relevant scientific community—distinguishes our position from that of instrumentalists (e.g., Friedman 1953), who regard prediction as a sufficient test of an explanation's adequacy. Reducing the time lag between cause and effect produces more detailed and fine-grained explanations and decreases the possibility

¹⁰ For Bunge (1996, p. 138), mechanisms are necessary to explain the behavior of all outcomes—save for those involving the behavior of things having no parts, such as elementary particles and photons. "As soon as we study systems, from atoms to nations, we experience the need to know how they work? . . . The only condition for a mechanism hypothesis to be taken seriously in modern science or technology is that it be concrete (rather than immaterial), lawful (rather than miraculous), and scrutable (rather than occult)." Note how much definitional leeway this view of mechanisms provides.

¹¹ Hedström and Swedberg (1996, p. 32) state that "intentional explanation allows us to 'understand' the act, in the Weberian sense of the term." Mechanisms need not be invoked at lower levels of analysis. Selection mechanisms used in neoclassical economics, evolutionary biology, and population ecology theory are examples of macrolevel causal mechanisms.

of spuriousness. Ultimately, the third criterion is the most useful one. Even if social mechanisms are unobservable, every distinctive mechanism is likely to have at least one unique empirical implication. In principle, then, rival mechanisms can be assessed by evaluating their unique empirical implications.

Somers (p. 770) rejects our criteria for evaluating causal mechanisms, arguing that they are necessarily tied to a particular substantive theory—rational choice—and thus not generally useful. She focuses primarily on one of our three criteria, reducing the time lag between cause and effect.¹² Somers contends that the emphasis on reducing the time lag between cause and effect, which we suggest is usually done by invoking more microlevel or mesolevel factors, is simply a substantive justification for rational choice theory masquerading as an epistemology. But rational choice theorists are far from unique in regarding mechanisms as means of more closely linking cause and effect by moving to lower levels of analysis. Stinchcombe (1991), one of Somers's favorite authorities, makes the same point using a variety of substantive examples. Moreover, Somers (pp. 770–71) unwittingly demonstrates that this criterion for judging mechanisms is not limited to rational choice theory, for she deploys it similarly in one of her own examples.¹³

Given our realist focus on causal mechanisms and abstract models, how can Somers (p. 746) accuse us of mixing realism and positivism into an incoherent brew? Even though our discussion of philosophical foundations is truncated, Somers's accusation of incoherence requires textual justification. Her strategy is to assume that any phrases or citations in the 1991 article that could possibly be linked to positivism provide proof positive that we are at heart positivists. She begins by labeling Hume and Hempel as two important founders of positivism and attempts to associate us with them. She (p. 747) refers to Hume as our "original authority" even though he is cited only twice and only two or three sentences are devoted to his ideas. She tries to link us several times to Hempel's deductive-nomo-

¹² It is hard to imagine that our other two criteria, plausibility and empirical implications, could in any way be tied to the substance of rational choice theory.

¹³ Somers argues that "since food has calories, and calories are energy, when we reduce our intake then the body has less energy to draw from external sources so it has to turn to internal sources of energy, which are stores of fat, and it uses up fat when it draws that energy" (p. 770). Note that her argument does exactly what we advocate—it reduces the time lag between the amount of food consumed and weight loss by filling in the microlevel processes linking the two factors. In this case, these microlevel factors ultimately derive from molecular biology. Thus her own example proves her argument wrong: our epistemological criteria are not tied to the substance of rational choice theory.

logical positivism, even though the only time we mention Hempel is to distinguish our position from his (Kiser and Hechter 1991, p. 6).¹⁴ Somers's construction of an internal contradiction in our position, therefore, arises out of selective misinterpretations of fragmentary textual evidence.¹⁵

Somers's own philosophical stance apparently owes much to recent relativistic critiques of science.¹⁶ Thus, she (p. 746) denies that most social scientists consider causality to be a necessary feature of an explanation,¹⁷ and that causal relations and mechanisms are required for adequate explanations (p. 746).¹⁸ She (p. 741) contends that all knowledge, logics, and reasoning practices are "history-laden," and that "science does not exist outside its own historical conditions" (pp. 738–39). For Somers, history determines the theories that are accepted as true and the facts that are deemed important at any given time. She proposes a "historical epistemology" based on the "thesis that the history of a thing (and not just the logic of its construction) can tell you something fundamental about its nature" (p. 731).¹⁹ Somers describes her new philosophy of science as "relational realism": the "thesis that belief in the causal power of unobservables—

¹⁴ She also labels us positivists by fiat. Somers claims that we use a "positivist language of deduction" (p. 753) and that we have a "positivist concern" (p. 747) for the problems caused by the inability to directly observe causal mechanisms. She argues that our use of the phrase "causal explanation works by subsuming events under causal laws" links us to Hempel's positivism, but that phrase is a quotation from Elster (1983, p. 26), who is anything but a positivist. Our concerns with deduction, with mitigating problems due to unobservability, and our desire to discover causal regularities are by no means confined to positivism—these issues are also central features of realism.

¹⁵ Such misattribution occurs frequently in Somers's article. For another example, she claims that "Kiser and Hechter now appear to follow Friedman's conception of 'positive economics'" (p. 759)—but we do not endorse Friedman's views and never have.

¹⁶ To a large extent, this is also true of Quadagno and Knapp (1992, p. 504). They note that "behind the idea of generating testable implications lies a presumption that science should seek general laws that are abstract representations of reality" and reject this conception of science. For a recent critique of scientific relativism, see Koertge (1997).

¹⁷ Although she claims that no such consensus exists, the only citations she provides are to Hume, Comte, and Popper—no contemporary sociologists are cited who disagree with our claim. It is true, as Quadagno and Knapp (1992, p. 487) point out, that some interpretive sociologists renounce any concern with causality. However, we still maintain that the vast majority of social scientists interested in explanation regard causality as central to the enterprise.

¹⁸ Presumably, she is alluding to the continued existence of some traditional positivists who do not believe in causal mechanisms. This was a popular position in sociology a few decades ago, but it has few adherents now.

¹⁹ One of the central claims in our 1991 article was that historical sociologists were increasingly replacing the sociological norms concerning research practice with the norms of historians. The philosophical foundations of Somers's position are a perfect illustration of this, as she notes: "relational realists borrow freely from what has traditionally been historians' terrain" (p. 767).

such as states, markets, or social classes—does not depend on the rationality or truth of any given theory but upon practical evidence of its causal impact on the relationships in which it is embedded” (p. 744). She goes on to outline three principles of relational realism (p. 766). The causal power of a “theoretical social dynamic” (her examples are gender and states) must be “independent from absolute belief in any one particular theory.” The theory must have an epistemology (in this case, her “historical epistemology”). Last, “there are no universally valid principles of logical reasoning; there are only problem-driven ones.” Somers’s final principle is the most telling testament to her extreme relativism; it implies that she considers even the most basic rules of logic to be culture bound. She claims that there are no universally valid principles for logical reasoning; this has the absurd implication that rules of logical inference such as *modus ponens* and *modus tollens* are valid in some historical periods but not in others. These statements harken back to a conventionalist philosophy of science that regards truth as nothing more than a historically determined social construction.

We are unpersuaded by such views (see Laudan 1990; Searle 1995). The degree to which any phenomenon is a social construction ought to be regarded as an empirical question rather than an epistemological axiom.

THEORY AND ITS ROLE IN EXPLANATION

These days there are almost as many definitions of theory as there are sociologists. Does this mean that the term is impossible to define or that any one definition of theory is as valid as any other? Not at all.²⁰ Definitions of theory have consequences for the quality of research, the ability to test arguments, and the cumulation of knowledge. The principal claim of our article was that general theory is useful in historical explanations, since it is an important source of causal mechanisms. General theories provide both the omnitemporal laws that animate contextual models (Popper 1994, p. 165) and the guidelines necessary to attack particular substantive problems.²¹ Our critics responded primarily by claiming that they do

²⁰ Here are two definitions of theory that we endorse. “[Theories] are *causal* explanations providing *intelligible* answers to why-questions about *empirical* facts” (Hechter 1987, p. 1, n. 1). “Theory can be taken to mean a set of assumptions or postulates with which one approaches some part of the empirical world, a set of concepts in terms of which this part of the world is described, and a set of propositions, emerging from the assumptions and relating the concepts, about the way this part of the world ‘works,’ which are checked against observations of that world” (Stryker 1959, p. 111).

²¹ General theories “supply the group with preferred or permissible analogies and metaphors. By doing so they help to determine what will be accepted as an explanation and as a puzzle-solution; conversely, they assist in the determination of the roster of unsolved puzzles and in the evaluation of the importance of each” (Kuhn 1970, p. 184).

not in fact reject theory; rather, they plead for a different type of theory. They advocate a looser definition of theory and suggest that it should be constructed historically and that its main components (such as mechanisms and scope conditions) should be defined historically instead of abstractly.

Much of the debate generated by our article focuses on the relative importance of induction and deduction in the conduct of research. Although we were critical of an overreliance on induction in current historical sociology,²² we never advocated an equally one-sided emphasis on deduction. In fact, we explicitly renounced a purely deductive methodology and praised research that combined induction and deduction (p. 17).²³ In spite of this, our critics insist on the caricature that we are pure deductivists (Quadagno and Knapp 1992, p. 491; Skocpol 1994, p. 322; Somers, in this issue, p. 725). But this conclusion rests on a misreading of our original text. We share with our critics the conviction that no research in comparative-historical sociology is purely inductive, just as none is purely deductive. All good research combines both deductive and inductive aspects.

The crux of the debate turns on the kinds of ideas and theories that are likely to stimulate productive research. We hold that research ought not be formulated merely on the basis of loose ideas or orienting concepts, but on the basis of general theories having explicit assumptions and abstract scope conditions. Our point was not that current comparative-historical sociology lacks ideas, but that it ignores the kinds of ideas derived from general theory.²⁴ General theories constitute paradigms that

They also provide researchers with shared exemplars that allow them to construe their problems as analogous to problems whose solutions have already been attained. The researcher "discovers . . . a way to see his problem as *like* a problem he has already encountered. Having seen the resemblance, grasped the analogy between two or more distinct problems, he can interrelate symbols and attach them to nature in the ways that have proved effective before. The law-sketch, say $f = ma$, has functioned as a tool, informing the student what similarities to look for, signaling the gestalt in which the situation is to be seen . . . He has . . . assimilated a time-tested and group-licensed way of seeing" (Kuhn 1970, p. 189).

²² Skocpol (1994, p. 322) argues that she started with some theoretical notions from Weber, Hintze, and resource mobilization theory, and thus claims that we incorrectly label her work as primarily inductive. However, the fact that scholars as diverse as Burawoy (1989) and Sewell (1996) also describe her approach as inductive lends credence to our interpretation.

²³ We explicitly commented that "induction is necessary in comparative-historical research" (1991, p. 24).

²⁴ This approach—beginning not with explicit general theory but only with some "theoretical notions" or "orienting concepts"—is common not just in historical sociology, but in the discipline as a whole. We consider it justified only when no good theory exists in a particular area. In that case all research must be exploratory.

can lead to research programs. They consist of explicit assumptions about causal relations, mechanisms, abstract models, and conditions delimiting the scope of such assumptions. Together these elements can be used to generate testable causal propositions. The assumptions, mechanisms, and models are part of the hard core of the theory (Lakatos 1978) that are not directly tested,²⁵ and are questioned only in the last instance. The utility of one general theory relative to others is judged in part by its logical coherence, but ultimately by how much it explains about relevant empirical observations.²⁶ A fruitful general theory can generate many propositions using relatively few basic assumptions, mechanisms, and models (Jasso 1988). The more empirical support these propositions receive, the more useful the theory.

In contrast, our critics seem to think that almost any type of idea qualifies as a theory. Quadagno and Knapp (1992, pp. 493, 504) suggest that historical sociologists need start only with some "initial theoretical notions," and that "theory should provide the researcher not with statements to be tested, but with questions."²⁷ They advocate a "temporally grounded historical sociology," that stresses contingency and sequentiality, and focuses on "the theoretical value of narrative."²⁸ If theory can refer to any "theoretical notions" (Quadagno and Knapp 1992, p. 493), "theoretical social dynamics" (Somers, in this issue, p. 766), or particular substantive arguments, then Quadagno and Knapp (1992, pp. 482–84) are surely cor-

²⁵ They are tested indirectly, by tests of the propositions they generate. None of these individual tests is conclusive, so theories are never abandoned until incorrect predictions (unresolved anomalies) cumulate and alternative theories arise that can explain them.

²⁶ Quadagno and Knapp (1992, p. 495) argue that using general theory leads to empirical problems in that "the risk of selecting confirming evidence while ignoring disconfirming evidence is greater when theory is defined as universal and omnitemporal." This would be true only if case choice procedures were totally unspecified, since the broader scope would allow scholars to engage in more successful ad hoc searches for confirming cases. However, one of the central arguments in our 1991 article was that cases should be carefully and systematically chosen in order to prevent this outcome.

²⁷ Quadagno and Knapp (1992, p. 504) argue for a more open-ended beginning to the research process and suggest that our approach is too closed. "In their metatheoretical program, theory structures historical inquiry in such a way that the 'ending point' of research is given prior to the conduct of research itself. The answers are known before the questions are asked." It is true that beginning with general theory structures inquiry more than beginning with some vague hunches or orienting concepts, but this only means that the questions to be asked are clear, not that the answers are already given. We stressed the fact that anomalies often arise in the process of doing research, and that these anomalies often modify explanations. Thus beginning with general theory does not mean that all of the "ending points" of the research are known in advance.

²⁸ For example, Quadagno and Knapp want to "derive theoretical explanations from detailed narrative histories" (1992, p. 482).

rect when they assert that most historical sociology is theoretical.²⁹ By the same token, so is the curriculum at McDonald's Hamburger University.

Although philosophers of science consider mechanisms to be abstract and omnitemporal (Bunge 1996), our critics prefer them to be historically specific.³⁰ Above, Somers (pp. 768–69) argues for “situational” mechanisms,³¹ and stresses the “contingent and indeterminate nature” of her “basic causal mechanism,” the “contingent relational pathway.”³² Quadagno and Knapp (1992, p. 486) argue along similar lines that narratives can identify and specify mechanisms.

Yet historically defined mechanisms are difficult to assess since—by definition—they do not apply to other situations or to substantive problems. Such mechanisms can have little utility for other researchers. In contrast, the mechanisms derived from general theories are *generalizable*—they can be used in different substantive areas and historical periods. Hence, even when exploring a historically unique set of causal relations, it is always preferable to invoke general mechanisms.

How should the scope of theory be defined for the purposes of historical

²⁹ Somers's substantive theory is a blend of network ideas (the relational part) with a social constructionist view of the micro level as consisting of “identities” that shape action. For her, action is based on “socially constructed identities” (p. 766). She argues that “identity can only be deciphered by virtue of its ‘place’ in relationship to other identities in its web.”

³⁰ Somers takes evolutionary biology as an example that supports her approach (pp. 771–72), but her reading of this field—which is heavily indebted to the writings of Steven Gould—is tendentious. Gould himself is embroiled in a debate about the proper role of theory in biology that in many respects parallels this one (to get some flavor of the debate, see Dawkins [1997] and Gould [1997]). For a radically different view of the state of evolutionary biology, see Sober (1993), and Dennett (1995). Somers argues that evolutionary biology “lacks true invariant laws” (p. 771), but this is a serious misunderstanding of the form of explanation characteristic of this discipline. Evolutionary biology illustrates well the necessity of combining general theory and initial conditions—i.e., historical details. Although path dependence and historical contingency are important parts of the narratives of evolutionary biologists, the general theories of natural selection and molecular biology undergird all such stories. The former provides macrolevel causal mechanisms, the latter microlevel causal mechanisms.

³¹ Somers cites Stinchcombe (1991) on situational mechanisms, but clearly he does not view mechanisms as contingent and indeterminate, as she claims; rather, he considers them to be general, stable features of recurring situations.

³² It is not exactly clear what this means, and the examples Somers provides do not enlighten. For example, she claims: “There are no laws for the extinction of species, for example, but there are indeed mechanisms: a comet hits . . . kills the dinosaurs thousands of years later” (p. 771). Somers is mistaken: the Alvarez theory of the extinction of dinosaurs posits the impact of an asteroid, not a comet. Comets are formed mostly of ice and could not have had the same catastrophic impact. Is this an example of a “contingent relational pathway”? We would consider the asteroid's impact to be an exogenous shock, not a causal mechanism.

analysis? On the face of it, scope conditions can be specified either historically or abstractly. Historically defined scope conditions (favored by Quadagno and Knapp [1992, p. 501] and by Somers) are based on particular spatial and temporal parameters (e.g., a given argument may only apply to 17th-century France). In contrast, abstract scope conditions merely contain general specifications of conditions that could exist in many places and times (e.g., a given argument may apply to hereditary monarchies).

Historical scope conditions can always be translated into abstract ones by redefining particular features of historical cases as values on a set of parameters.³³ This translation makes arguments more explicit, more general, and more testable. Perhaps the most important virtue of using abstract scope conditions is that arguments developed in one place and time can be applied to others. As in the case of general theory, the use of abstract scope conditions increases the generalizability of explanations. When scope conditions are defined abstractly, the number of relevant cases of the given phenomenon is expanded, making it easier to test arguments. Since historical sociologists often are plagued by having a surfeit of variables relative to cases, this is an especially important advantage.³⁴ Finally, abstract scope conditions increase the relevance of historical sociology by allowing insights developed in the study of past societies to be applied to present and future states of the world. This is the most important contribution that historical sociology can make to the discipline as a whole.

RATIONAL CHOICE THEORY AND THE EXPLANATION OF HISTORICAL OUTCOMES

The primary purpose of our initial article was to argue for the use of general theory in historical research. Since rational choice was our only extended example of a general theory,³⁵ much of the ensuing debate has

³³ In most cases, the use of historical instead of abstract scope conditions is the result of a failure to fully explore the implications of an argument—in other words, of an incomplete theory. Saying that an argument applies to 17th-century France, e.g., is a vague shorthand for a set of characteristics affecting the causal relationship in question that presumably are present in France in the 17th century but absent elsewhere.

³⁴ Historically defined scope conditions are unproblematic in noncomparative history. However, comparative historians must rely on abstractly defined scope conditions to select their cases.

³⁵ Our use of a sociological version of rational choice theory to address historical issues is by no means novel. A theoretical and methodological approach to comparative and historical sociology that uses clear abstract models and intentional action as its main causal mechanism already has been articulated, although it has neither been systematically explicated nor widely emulated. This is the approach developed by Max Weber in part 1 of *Economy and Society* ([1922] 1968). Below, we sketch what we construe

challenged its merits. Most of the faults our critics find in rational choice, however, derive from timeworn critiques of neoclassical economics. These critiques are not relevant to the much different form that the theory has assumed in recent sociology (for a review, see Hechter and Kanazawa [1997]). The writings of sociologists using rational choice, including Julia Adams (1996), Terry Boswell (Boswell and Dixon 1993), Mary Brinton (1988), William Brustein (1996), Ivan Ermakoff (1997), Debra Friedman (1995), and Diego Gambetta (1993), are not easily confused with those of George Stigler and Milton Friedman.

Our critics have different kinds of reservations about rational choice theory. Somers not only rejects rational choice, but all forms of methodological individualism and intentional explanation, as well. She rejects the contention that purposive agents determine social outcomes (pp. 763–64) and denies that complete explanations must include discussions of the actions of individuals. Her most fundamental criticism of rational choice theory is that it relies on an implausible ontology (pp. 762–66): the microfoundations of rational choice theory are too simplistic since people do not always act rationally. This charge damns all general theories, not just rational choice. Since all theories simplify reality, their ontologies are necessarily incomplete, but this does not make them implausible.⁶ Finally,

as a neo-Weberian approach to historical sociology, although the philosophical foundations of our approach cannot be traced to Weber, who was a nominalist rather than a realist. Weber was a methodological individualist (1968, pp. 4,7,13) who argued that it was preferable to begin analyses by assuming instrumental motivations (p. 5) and using abstract models of forms of social relations (ideal-types; pp. 19, 21; for a pioneering attempt to analyze an historical outcome that relies on both Weber and rational choice theory, see Hall [1988]). Like earlier readings of Weber by Parsons (1937), Bendix (1960), and Collins (1986), our interpretation is shaped by our own theoretical commitments. We believe that the core of Weber's approach to historical analysis corresponds to emerging developments in sociological versions of rational choice theory (Hechter 1987; Brinton 1988; Levi 1988; Coleman 1990; Kiser 1994; Adams 1996; Brustein 1996; Hechter and Kanazawa 1997). Our rational choice interpretation of Weber is roughly analogous to interpretations of Marx proposed recently by Elster (1985) and Roemer (1986). Just as they outlined an "analytical Marxism," here we develop an "analytical Weberianism." However, our arguments must stand or fall on their own merits, not on the legitimacy of Weber. For this reason we do not emphasize our debt to Weber in the following discussion.

⁶ This criticism is partly a result of her confusion about the microfoundations of rational choice theory. Somers's definition of rational choice theory conflates its core element (rationality) with various auxiliary features (self-interest, wealth maximization) without distinguishing the two. She claims that "the basic ontological assumptions of rational choice are well-known—that all actors apply the standards of means-ends rationality, that they are self-interested, and they are wealth maximizers" (p. 763). The hard core of rational choice includes only the first of these assumptions—action is purposive, intentional, and involves means-ends calculations. The other two, self-interest and wealth maximization, are common auxiliary assumptions used in many, but far from all, applications of rational choice (Friedman and Diem 1993). Rational

Somers claims that models used in rational choice theory cannot deal with temporality and path dependence,³⁷ and cannot produce narrative analyses that incorporate particular historical details—thus they are inadequate tools for analyzing history (cf. Kiser 1996).³⁸

In contrast, Skocpol sees nothing amiss with rational choice micro-foundations on theoretical grounds and even admits that she has been using them (implicitly) herself.³⁹ Skocpol's principal complaints are substantive and empirical. She regards rational choice arguments simply as post hoc redescriptions of findings already discovered by others. For Skocpol (1994, p. 325), the jury is still out on whether rational choice theory can improve on existing explanations of macrosociological outcomes.

These charges deserve a response. To that end, we next outline the rational choice research program in historical sociology and document some of its substantive successes.

The Role of Actors in Historical Sociology

How should sociologists explain social outcomes and events? We contend that all good sociological explanations must consist of separate arguments

choice theorists have been increasingly interested in noninstrumental motivations recently, including those based on values (Hechter 1992a, 1994; Friedman, Hechter, and Kanazawa 1994), emotions (Frank 1988), and habits (Becker 1996). Admittedly, most research inspired by rational choice in historical sociology has failed to incorporate noninstrumental motivations.

³⁷ Somers (pp. 754–55) argues that rational choice theorists advocate “minimal temporality,” but the quotation she uses to make that attribution is taken entirely out of context. Our discussion of “minimal temporality” was part of an argument about the virtues of minimizing the time lag between cause and effect—it had nothing to do with any rejection of the importance of sequence or path dependence.

³⁸ Somers's misunderstandings of rational choice deepen when she delves into substantive arguments. For example, she asks “we can count low voter turnout, but how do we ‘prove’ that the cause is really utility maximization and/or free ridership, as opposed, say, to production of registration difficulties put in the way of voting by state and county governments?” (p. 740). But the latter is not an argument opposed to rational choice, it is a central part of it—a rational choice model predicts a strong association between the costs of registration and voter turnout. A similar difficulty arises in her discussion of Friedman (1953). She claims that Friedman's argument about the relationship between markets and firms lacks a causal mechanism because it does not rely on the thought processes that produce managers' choices. She is right that the actual choice process of managers does not matter to Friedman, but the reason it does not is that Friedman relies on another powerful causal mechanism, market selection. Regardless of what managers think or do, he argues that markets will select the most efficient firms. His mechanism is not invoked at the micro level.

³⁹ “The reason that Taylor is so easily able to provide a rational-choice gloss on my discussion is that he and I were starting from much more similar metatheoretical assumptions than he supposes” (1992, p. 325). Skocpol elaborates, “I simply assumed that peasants and others whose actions I narrated were acting rationally” (p. 325).

pertaining, first, to the *motives* of individual actors and, second, to *models* of the contexts within which their action takes place. Good explanations in historical sociology are no different, save in one crucial respect. It is far more difficult to explain outcomes and events in societies distant in time and space than in our own.

All complete social explanations must include an analysis of individual motives and actions (Hechter 1983; Coleman 1986). For sociological purposes, individuals are the elementary units of analysis, for they cannot be divided into smaller acting units. If outcomes and events ultimately are the products of the actions of individuals, just how are these phenomena to be explained? What accounts for individual action? Since understanding the reasons why people act as they do is essential to explaining outcomes and events, we must employ a different methodology than that used in the natural sciences.

Motives, for our purposes, may be usefully divided into two elements: *orientations to* and *goals of* action. Whereas goals are practically infinite in their variety, orientations to action are either instrumental or not. Instrumental action is consequentialist. People are motivated instrumentally whenever they choose a course of action that they believe is the most effective means of attaining their goals in a given situation. Instrumental action implies nothing at all about the nature of these goals, however: they can be materialist or idealist, conventional or wildly idiosyncratic. The wealth-maximizing capitalist and the salvation-seeking cleric are both instrumentally motivated.

Sometimes, however, orientations to action are noninstrumental. For instance, action may be determined by a conscious belief that a given ethical, aesthetic, or religious behavior must be taken because of duty, right, or merely for its own sake, regardless of its consequences.⁴⁰ This is value-oriented action. Emotional action, which is determined by the actor's specific affects and feeling states, is likewise nonconsequentialist. Emotional action is carried out in the heat of the moment and without forethought—"crimes of passion" are a classic example. Finally, action can be purely habitual—determined by reflex rather than calculation or impulse.

Do instrumental orientations to action provide the best basis for explaining social outcomes and events? This is an empirical question. Often, however, direct and reliable evidence about orientations to action is difficult to come by. When, due to a lack of evidence, it is necessary to make

⁴⁰ Parsons (1937) regards ritual as an exemplary form of value-oriented action. This, however, is far from certain. Insofar as rituals are performed in public settings, participants will have instrumental reasons for engaging in them.

assumptions about action orientations, explanations should proceed by assuming that actors are instrumentally oriented. Instrumental action is least ambiguous and therefore most understandable to the analyst, who may not share an emic perspective with her subjects. Hence, by assuming an instrumental orientation, the analyst can most easily derive models that yield empirical implications for any set of circumstances. Substantively, how could individuals or cultures whose action is predominantly nonconsequentialist ever successfully compete against instrumental actors?

Since instrumental action is intentional rather than random or unconscious, it implies goals. If we do not understand the actor's goals, then we cannot deduce how an instrumentally motivated person would go about realizing them. To understand an action we must know the goals people hope to attain by undertaking that particular action and the information available to them, including their beliefs. Last, we must appreciate the contextual circumstances they face that limit or enhance their ability to attain these goals.

Whereas actual behavior may be public, the mechanisms generating it (intentional action based on goals) are internal states that can only be seen as through a glass darkly (Hechter 1992*b*; Hechter, Nadel, and Michod 1993). Since an instrumental explanation of outcomes and events requires specification of actors' goals, the difficulty of discerning these goals often mitigates our attempts at explaining outcomes.

Goals can be imputed at different levels of abstraction, and on this account with more or less realism. Most realistically, we can attempt to ascertain the actual goal in the given case of a particular actor or the average and approximate goal attributable to a given group of actors. Alternatively, we can assume that hypothetical actors are motivated to pursue given goals on the basis of theoretical considerations. Clearly, we would always prefer to employ the most realistic motivations in our explanations, but our ability to do so is limited by the availability of appropriate evidence. The greater the availability of evidence, the less we need to employ motivational assumptions to explain social outcomes. For example, to explain contemporary events we can exploit our knowledge of the relevant circumstances that we share with the actors in question as members of the same society. Even so, it is unlikely that we will be very knowledgeable about the precise nature of the actors' goals and circumstances. Since contemporary actors are, in principle, available for us to study, we can employ a battery of different data collection techniques—ranging from sample surveys, to intensive interviews, to participant observation methods—to elicit their goals and circumstances.

Unfortunately, however, we usually know much less about the goals

and circumstances of individuals in societies that are distant in time or space than we do of those closer at hand.⁴¹ On this account, historical sociologists must place greater reliance on motivational *assumptions* in their explanations. Therefore they must be sophisticated about deriving models from assumptions about goals and instrumental action.

Since the number of possible goals is astronomical, what kinds of assumptions should be made? We recommend that analysts begin by assuming individual goals such as the pursuit of wealth, power, and prestige. These particular goals are fungible—they can serve as means to the attainment of a plethora of individual ends. Unlike nonfungible goals, which are more likely to be idiosyncratically distributed in any society, everyone can be relied on to prefer outcomes that promise them greater wealth, power, and prestige to those promising less (Hechter 1994). Further, these goals also have straightforward behavioral implications, making it feasible to construct testable models.

Sociologists inspired by rational choice have been able to construct precise and detailed explanations using these simple microfoundations. A notable example is William Brustein's (1996) *The Logic of Evil*. In sharp contrast to the prevailing wisdom that the rise of the Nazi party was due primarily to anti-Semitism, Brustein begins with the simple assumption that Germans chose to affiliate with the party whose platform best served their material interests. Brustein analyzed political party platforms in detail and collected "the most systematic and comprehensive database of Nazi membership files that exists today" (Anheier 1997, p. 200) to test his arguments. The propositions generated by his theory and supported by the historical data are precise: he explains why livestock workers, import-oriented blue-collar workers, and skilled laborers joined the Nazis much more frequently than grain growers, export-oriented blue-collar workers, and unskilled laborers. Clearly this is no case of post hoc redescription of findings that others have already discovered, but a systematic analysis guided by general theory that produced a novel and complex understanding of an important historical event.

Models of Context in Historical Analysis

A common misunderstanding of rational choice is that it is a microlevel theory. Somers (p. 763) succumbs to it by defining rational choice only in

⁴¹ Weber (1968, pp. 5–6) notes that "many ultimate ends or values toward which experience shows that human action may be oriented, often cannot be understood completely, though sometimes we are able to grasp them intellectually. The more radically they differ from our own ultimate values, however, the more difficult it is for us to understand them empathetically." Tilly also comments on the "difficulty of verifying arguments emphasizing attitudes and intentions with this sort of [historical] documentation" (1975, p. 9).

terms of its microlevel assumptions. She ignores the macrolevel models of rational choice, even though they were a major part of our 1991 article. Rational choice is a multilevel theory—it focuses on explaining macrolevel outcomes, and it does so with arguments that always combine macro and micro levels (Coleman 1986, 1990; Friedman and Hechter 1988; Hechter and Kanazawa 1997). Thus in addition to assumptions about motives and goals, we need a method for analyzing the contexts (including the structural and cultural conditions) within which action takes place.⁴²

As in the case of motives, contexts can be analyzed either in concrete empirical terms (attempting to specify the particular constraints faced by actors in all of their historical detail) or in relation to abstract theoretical models of major categories of such contexts. The latter choice is preferable on several counts.⁴³ As in the case of motives, abstract models of contexts serve to reveal anomalies and to foster the cumulation of knowledge. They permit comparisons across cases (even those in very different places and times) and across substantive topics. Most fundamental, they are useful in generating testable hypotheses. One of the major advantages of abstract models over the use of classification schemes or the orienting concepts advocated by our critics is that they specify the interrelations between factors. As Rodney Stark (1997, p. 26) puts it, "The difference here is that between a parts catalog and a working diagram of an engine . . . a model explains why and how things fit together and function."

What sorts of contextual models should we construct? Since we assume an instrumental orientation to action and egoistic motivations at the micro level, the answer is models that are consistent with these assumptions. The building of such models is one of the principal activities of rational choice theorists, who have developed a number of endogenous models of context from agency theory, game theory, and group solidarity theory (see Schelling [1960], Becker [1976], and Friedman and Hechter [1988] for detailed discussions of several of the models developed in rational choice theory). However, the analysis need not be limited to models developed on the basis of rational choice. Several other types of models are consistent with these microfoundations, including models of economic structure and group conflict derived from Marxism (see Elster 1985; Roemer 1986; Przeworski 1991), models of social relations from network theory (Cole-

⁴² Some historical sociologists are concerned with the origins of particular historical contexts within which instrumental action takes place (e.g., Carruthers 1996; Clemens 1997).

⁴³ Strictly speaking, the first choice is impossible—to describe fully the context of an action would take longer than the life expectancy of most scholars. All of us have to choose to focus on some elements of the context to the exclusion of others.

man 1990; Burt 1992), and of course Weberian ideal-types of organizational forms and political systems (Weber [1922] 1968). For example, Boswell and Dixon's (1993) analysis of the relationship between exploitation and rebellion begins by reconstructing and elaborating Marx's theory, using ideas drawn from analytical Marxism, and then provides a systematic test of that theory based on quantitative cross-national data. Adams (1996) combines the use of agency theory with Weber's model of patrimonial states and a model of brokerage from network theory to construct a fascinating analysis of the relationship between states and colonial trading companies in England and the Netherlands. Hedström (1994) combines rational choice with network analysis to explain the spatial diffusion of labor unions in Sweden from 1890 to 1940. These syntheses of rational choice microfoundations and compatible models taken from other sources illustrate the influence of rational choice in historical sociology—and how different it is from its neoclassical cousin.

The abstract scope conditions of rational choice theory render these explanations generalizable; this, in turn, increases their testability and contemporary relevance. For example, Stark's (1997) use of general theory in his analysis of the rise of Christianity allows him to compare the conditions that facilitated its rapid growth to the contemporary case of the Mormons. Moreover, Kiser (1994) uses agency theory to argue that, when communications, transportation, and record-keeping technologies are inadequate, particular types of monitoring problems will arise in the collection of indirect taxes that are best mitigated by using tax farming (privatized tax collection) instead of salaried state administrators. Although the argument was first tested in the context of early modern states, since its scope was abstract—applying to conditions in which technologies of communication, transportation, and record keeping are poor—Kiser and Baker (1994) were able to use it to explain forms of tax administration in contemporary less developed countries.

Rational choice theorists have often focused on ways that temporal sequence in which causal factors occur affects outcomes.⁴⁴ One set of rational choice models that can incorporate temporality come from game theory, especially extensive form games (Kreps 1990, p. 13). These models explain outcomes as the consequence of temporally ordered strategic inter-

⁴⁴ Temporality is also important at a more micro level. Rational choice research on discount rates (the rate at which actors discount future costs and benefits relative to current ones) has contributed to our understanding of this process. Most economists assume that discount rates are normally distributed in the populations they study, a reasonable assumption if one does not know what causes discount rates to vary. More sociological rational choice models have been attempting to discover the structural determinants of variations in discount rates. For example, Levi (1988) shows that discount rates increase with the insecurity of rule.

action—sequences of action and reaction. Abell's (1987) integration of game theory into narrative analysis is the most highly developed attempt to incorporate sequentiality in a theoretical manner, but there are other prominent examples. Lindenberg (1989) outlines an ordered sequence of game structures that are likely to unfold as a revolutionary situation moves toward revolution and then employs these same structures to construct brief narratives of the French and Russian revolutions. Heckathorn (1988, 1990) uses game theory to demonstrate the importance of collective sanctions in maintaining social control within groups. Drawing on Heckathorn, Brown and Boswell (1995) analyze past strike outcomes by combining narrative, game theory, and qualitative comparative analysis (QCA) to demonstrate that the sequence of "moves" by workers and unions is a critical determinant of the outcome.⁴⁵

Of course, if our purpose is to explain unique social outcomes or events, abstract models and assumptions about motivations are insufficient. The explanation of a unique outcome requires both general models and initial conditions—that is, particular features of the specific case. However, these initial conditions need not be *ad hoc*. Good models suggest the specific features of cases that should be causally important. More fundamentally, abstract ideal-types allow for the separation of the general and particular features of each case, facilitating comparative analysis.

One feature of recent applications of game theory that increases its compatibility with historical sociology is the incorporation of the effects of institutions and even particular events. Formal game theory models often do not produce precise predictions of outcomes. The models yield multiple equilibria and cannot predict which of these will occur in a given situation (Hechter 1992a). At their best, these formal models reveal the outcomes that cannot occur—those that are off the equilibrium path. To make their explanations more precise, game theorists increasingly have attempted to integrate formal game theoretic models with particular features of the historical cases they study since only these particular features can tell them why one of many possible equilibrium outcomes in fact occurred. The

⁴⁵ Game theory is not the only means of incorporating temporality into rational choice explanations. Brinton (1988) shows that differences in the timing of human capital investments in Japan and the United States result in different levels of gender stratification. Barzel and Kiser (1997) demonstrate that the timing of factors affecting the insecurity of rule determines their effects on voting institutions—the Hundred Years' War disrupted the development of voting institutions in France more than in England because their prior development (and thus the ability to withstand shocks) was greater in the latter. Rational choice work on agenda setting also incorporates temporality, generally by demonstrating how moving first allows certain actors to shape outcomes (see Weingast and Moran [1983] on congressional committees and Mueller [1989, p. 255] on bureaus proposing budgets).

classic example of this strategy is found in Schelling's (1960) discussion of focal points.

More recently, game theorists have been integrating formal rational choice models with detailed narrative stories that incorporate initial conditions (Abell 1987; Bueno de Mesquita 1996; Weingast 1996; Bates et al. 1998). Ermakoff's (1997) analysis of the shift in medieval marriage norms is a case in point. Using Levi's (1988) model of quasi-voluntary compliance, a game theoretic model of bargaining between aristocrats and Roman prelates, and a detailed analysis of the historical evidence, he shows why late 11th-century nobles decided to obey church norms about divorce and endogamy (marrying cousins). These models provide a way of constructing explanations that combine the general and the particular while maintaining a clear distinction between the two.

Rational choice is also helpful in studying path dependence. This concept initially was developed by economists using rational choice micro-foundations (Arthur 1988, 1989; David 1985). It emerged as an attempt to resolve apparent anomalies in neoclassical economic theories—the persistence of ostensibly inefficient equilibria, such as the adoption of the QWERTY typewriter keyboard and the VHS videotape format.⁴⁶ The concept of path dependence combines specificity (the importance of particular initial conditions) and generality (abstract causal mechanisms that make it difficult to exit from particular paths). Both history and general theory are necessary to explain path-dependent outcomes, and neither alone is sufficient (Blute 1997). The point holds both for evolutionary biology and for historical sociology.

Rational choice theory is especially important in explaining the conditions under which path dependence will be strong or weak. For example, although the QWERTY keyboard configuration has remained constant, other aspects of typewriting, such as the shift to electric typewriters and then computers, have changed dramatically (Williamson 1993, p. 143). General theory is necessary to explain why some aspects of typewriter technology seem to be strongly path dependent and others do not (North 1991, pp. 92–95). Path dependence arguments that lack general theory, such as those that Somers advocates, lack sufficient causal mechanisms: they often take a *post hoc ergo propter hoc* form in which continuity over time is posited as proof of causal path dependence. Baron and Hannan (1994, pp. 1141–42) argue that “by articulating the sources of resistance to change in social systems, sociologists could help transform the study of

⁴⁶ It is not, however, evident that these two outcomes really are anomalous for neoclassical economic theory, since the efficiency differences between alternative formats for both products may have been overstated (Liebowitz and Margolis 1990, 1995).

path-dependent development from historicism into a field in which prediction and comparative assessments are possible." Rational choice theory can help do this in two ways. First, Arthur (1988, p. 10) has outlined four self-reinforcing mechanisms, all of which assume rational action, that produce and maintain path dependence: large fixed costs, learning effects, coordination effects, and adaptive expectations. Second, by providing abstract models of the payoffs to relevant actors of various strategies, game theorists can show why an outcome is path dependent—either because it is not in the interests of any of the relevant actors to change it or because actors lack the resources to do so. Path dependence is the result of low payoffs to actions that would lead to change—actions that are off the equilibrium path.

Testing Rational Choice Theory and Resolving Anomalies

Since the ultimate criterion of the value of any theory is its ability to explain the empirical world, our 1991 article focused on the importance of systematically testing theories.⁴⁷ Our critics disagree with our discussion of theory testing, but here too their interpretations of our position vary, and hence they offer quite different misgivings.⁴⁸ Quadagno and Knapp (1992, p. 494) note that our article "explicitly defines the process of identifying causal mechanisms as an empirical one [since] testing . . . involves

⁴⁷ There is ample discussion in our 1991 article of the importance of empirical testing (pp. 8–9), the importance of theories having multiple clear empirical implications (pp. 16–17); we used an empirical illustration (pp. 17–23) to show how our approach to state autonomy and policy formation differed from that of Skocpol (1979) and Mann (1984, 1986).

⁴⁸ Our arguments about the criteria for choosing cases to be studied also provoked responses. Somers argues that we "presuppose that cases are . . . given in the nature of things . . . rather than constructed along analytic parameters" (pp. 757–58). This is a very odd misreading of our position—How could supporters of theory-driven research advocate such an empiricist approach to case choice? We argued that cases should be selected and defined on the basis of theoretical criteria. Skocpol (1992, p. 332) defends *States and Social Revolutions* (1979) by asserting that it compares not countries but "episodes," implying that our criticism of the lack of independence of her cases is off the mark. However, episodes also must be independent in the Millian method: does Skocpol really want to claim that the 1917 "episode" of successful revolution in Russia was independent from the 1905 "episode" of failed revolution? Skocpol (1992, p. 323) admits that the cases she addressed were in no way a representative sample, but that her work should be judged in the broader context of subsequent studies that expanded the empirical base by studying other cases. It is certainly true that all research must be judged in the context of related research, and if many studies together produce a more complete (even exhaustive) sample, then the case selection problems in any particular study will be mitigated.

the search for historical data that will either verify or refute the proposed causal argument." They go on to (correctly) conclude that "Kiser and Hechter depend just as heavily on evidence for theorizing as do those whose work they disparage." In contrast, Somers drastically distorts our position by asserting that we do not care about history or the empirical testing of ideas. She complains that we have "an epistemology that tells us it is theory that adjudicates truth," (p. 755) and that ours is "a deductivism in which hypotheses are not only 'guided' . . . but adjudicated on the basis of general theory" (p. 762).⁴⁹ Somers suggests further that we "denounce the use of history" (p. 723) and that history is "a casualty of the new theory centrism" (p. 730).⁵⁰

None of our critics have much faith in the type of systematic empirical testing that we advocated. Quadagno and Knapp advocate a much looser version of testing explanations, implying that it is easy to test practically any claim. They suggest that "all forms of discourse are open to refutation, revision, or rejection, and every piece of historical research is "tested," regardless of the particular theoretical form in which it is constructed" (1992, p. 495). Somers takes the opposite line of attack; she lists several of the well-known problems with theory testing, leaving the impression that, for all practical purposes, the task is an impossible one (pp. 756–57).

Our own view lies in between these two poles. The testability of an explanation depends on the nature of available evidence and on the characteristics of the theoretical ideas. In historical research, evidence is often sparse and fragmentary. In such circumstances, only certain kinds of theories can be testable. To be testable, a theory must yield propositions.⁵¹ Further, the more general and fruitful the theory (the greater the number

⁴⁹ This sort of dismissive rhetoric is scattered throughout Somers's discussion. She describes our position as "theory proposes, theory disposes," and states that we have no "concern for accuracy and empirical detail" (pp. 730, 762).

⁵⁰ How could Somers misconstrue the intent of our arguments about the importance of empirical testing so thoroughly? This is accomplished by misinterpreting our philosophy of science and our definition of general theory. Somers downplays our commitment to empirical testing of ideas by *redefining* our philosophy of science as what she labels *theoretical* realism, which she claims is "the thesis that belief in an explanation depends on belief in the a priori theory from which it is imputed" (p. 727). Somers also equates general theory with what Lakatos (1978) calls the hard core of a research program, which is never directly tested (p. 743). However, Lakatos's hard core refers only to the fundamental assumptions of a theory. General theory also includes propositions derived from these assumptions, and these should be empirically tested.

⁵¹ This is the problem with Quadagno and Knapp's argument that the role of theory should not be to provide "answers" (presumably in the form of testable propositions) but only to raise questions for narrative analysis to address (1992, p. 505; see also Griffin 1993, p. 422). Since questions cannot be tested, they cannot generate anomalies, and most philosophers of science agree that discovering anomalies is the main source of cumulation of knowledge (Kuhn 1962; Lakatos 1978).

and diversity of its empirical implications), the easier it is to test.⁵² General theory entails all of these features, whereas none of them can be found in the alternatives advocated by our critics.

Whether a test confirms the empirical implications of theoretical models built on the assumption of instrumental action is always an open question. If the evidence is roughly consistent with the model's expectations, then the analyst may presume to have explained the outcome or event, especially if alternative explanations are less consistent with empirical observations. If the explanations are at variance with empirical findings, however, then the search for an adequate explanation must continue. How should the resulting anomalies be resolved? Anomalous findings can be due either to the analyst's misspecification of the context in which action occurs, to misperception of the actors' goals, to misspecification of action orientations, or to the inadequacy of the theory. We suggest that modifications of the explanation should be made first in the least central aspects of the theory, with consideration of successively more central elements if that tack fails. Thus, if a given model comes up short, first change the model of social structural constraints. For example, if group solidarity theory is used, have any types of monitoring or sanctioning been neglected? Perhaps an inappropriate model has been used to explain the phenomenon. Perhaps the case is really a repeated game as opposed to a one-shot game, or maybe it is a chicken game instead of a prisoner's dilemma. Only if all such modifications fail to improve the fit should one reassess the microfoundational assumptions (see Popper 1994, p. 17).

Consider the specification of the actors' goals. As we have seen, this is one of the most difficult things to determine and, no doubt, is the source of many of the incorrect predictions in historical explanations. Finally, if all else fails, perhaps the assumption of instrumental action is incorrect and some other form of action should be employed in the explanation. The outcome of this process of successive, staged revision invariably improves the models and better specifies their scope, thereby advancing our understanding of the given event or social outcome.

An example may help illustrate the way in which rational choice theo-

⁵² Somers favors a "problem-driven" approach to the choice of topics and cases instead of our theory-driven perspective (pp. 772-73). Basically this means that the choice of what to study comes from the realms of culture and current events rather than from theory. Somers seems to construe problems as *social* problems that arise in particular historical periods, often due to particular highly publicized events. She argues that history should shape the questions that guide research (pp. 733-34); thus, the revolutions of 1989 have increased scholarly interest in the study of democratization. No doubt, major historical events and serious social problems should and do affect the topics scholars choose to study. But this should not be the exclusive reason for choosing a research topic.

rists deal with anomalies. Almost all early modern states used tax farmers to collect indirect taxes in order to compensate for their poor monitoring capacity (Kiser 1994). However, Prussia—which likewise had poor monitoring capacity—employed state officials on fixed salaries instead. This would appear to leave Prussia highly vulnerable to corruption, but there is wide agreement that Prussian tax collection was efficient. In order to resolve this apparent anomaly, Kiser and Schneider (1994) explore how agency relations differed in Prussia. Agency theory suggests that either the monitoring capacity of Prussian rulers was superior or that there was another means of dissuading agents from corruption. It turns out that Prussia indeed had a unique way of collecting indirect taxes. It relied heavily on a “workfare” program that employed injured war veterans. (In other countries, injured war veterans were placed in hospitals or given welfare). Since injuries limited their alternative employment options, these agents were highly dependent on rulers for a job; being fired thus was a stronger sanction for them than it would be for uninjured agents. This high dependence inhibited corruption. Thus the apparent anomaly is resolved: Prussia could avoid tax farming because its workfare program provided state officials with an alternative means of overcoming their monitoring problems. Far from simply redescribing what everyone already knew, this study suggests that Weber (1968, p. 974) and many contemporary historians (Rosenberg 1958) may have been mistaken to conclude that Prussia’s efficiency owed to its bureaucratization: far from hiring tax collectors on the basis of merit, Prussia instead chose them on the basis of personal characteristics quite unrelated to their capacity to do the job. Indeed, deviations from bureaucracy may have made the Prussian state more effective than its competitors.

CONCLUSION

Debates on historical methodology have been taking place for ages. Our 1991 article began with a comment on the *Methodenstreit* that pitted German social scientists against historians in the early years of the century. If anything, the disciplinary scope of these debates has widened in recent years. Conflicts once confined to the human sciences now encompass critiques of the natural and physical sciences as well (Lynch and Bogen [1997] offer a recent programmatic statement). The thrust of Somers’s critique rests on a rejection of our epistemology. But epistemological differences are notoriously difficult to resolve. The advocates of rival views do not share enough common ground (the noncontractual basis of contract?) to engage in reasoned debate. Shouting matches are the usual result. We have no desire to engage in a shouting match. Suffice it to say

that we presume a standard form of realism that undergirds, at least implicitly, most empirical research in the social sciences. In contrast, Somers argues for a new, idiosyncratic epistemology that has yet to withstand the critical scrutiny of philosophers (see Searle [1993, p. 7] for a comment on how the vast majority of American philosophy departments do not endorse views of the sort that Somers espouses). Caveat emptor.

The debate over historical methodology will not be resolved by philosophers, nor should it be. Presumably, the point of the debate is to improve the quality of research in historical sociology. To that end, we have emphasized the research practices that flow from our approach (this debate would have been more productive had Somers done likewise). At the end of the day, historical sociologists must put their own chips down on a research strategy.

Why then is an approach based on general theory to be preferred? For at least five reasons. If all historical analyses dwell on some facts at the expense of others, what guides this selection process? General theories suggest which facts ought to be relevant and which can be safely ignored. Without explicit theoretical guidelines, implicit biases are likely to creep into the set of facts deemed relevant. Precisely because all observation is theory laden, the pretense to pure inductivism is at best self-delusory. Inductivism therefore relies on implicit rather than explicit theories, but surely social science is better off with explicit rather than implicit theories.

Second, such theories offer a source of generalizable causal mechanisms that suggest how and why given causes produce given effects. Third, general theories facilitate the cumulation of knowledge across substantive domains. Because their models and mechanisms are generalizable, within abstract scope conditions, this increases links between substantive domains as well as between the past and the present. Fourth, general theories are useful precisely because they do not provide perfect explanations. The anomalies they highlight are major sources of new ideas for research.

Last—and most important—research based on general theories provides the seeds of these theories' ultimate destruction. We advocate a sociological version of rational choice not because we hold the theory to be true or beautiful; we merely regard it as the best general theory currently on offer (see Abell 1992). We will be surprised and disappointed if it is possible to make the same assessment a few decades hence.

Therefore, we hope at some point to join Somers in saying, "Down with rational choice theory!" However, we disagree about the conditions of its eventual demise. She seems to think that rational choice will go away if it is shunned. Quite the contrary. Rational choice will disappear only when it is supplanted by a superior general theory. When that new day

dawns and the superior theory stands revealed, we strongly doubt that it will resemble the vision offered by Somers.

REFERENCES

- Abell, Peter. 1987. *The Syntax of Social Life*. Oxford: Clarendon Press.
- . 1992. "Is Rational Choice Theory a Rational Choice of Theory?" Pp. 82–94 in *Rational Choice Theory: Advocacy and Critique*, edited by James Coleman and Thomas Fararo. Newbury Park, Calif.: Sage
- Adams, Julia. 1996. "Principals and Agents, Colonialists and Company Men: The Decay of Colonial Control in the Dutch East Indies." *American Sociological Review* 61 (1): 12–28.
- Anheier, Helmut. 1997. "Studying the Nazi Party: 'Clean Models' versus 'Dirty Hands.'" *American Journal of Sociology* 103 (1): 199–221.
- Arthur, Brian. 1988. "Self-Reinforcing Mechanisms in Economics." In *The Economy as an Evolving Complex System*, edited by P. Anderson, K. Arrow, and D. Pines. Reading, Mass.: Addison-Wesley.
- . 1989. "Competing Technologies, Increasing Returns, and Lock In by Historical Events." *Economic Journal* 99:116–31.
- Baron, James N., and Michael Hannan. 1994. "The Impact of Economics on Contemporary Sociology." *Journal of Economic Literature* 32:1131–52.
- Barzel, Yoram, and Edgar Kiser. 1997. "The Development and Decline of Medieval Voting Institutions: A Comparison of England and France." *Economic Inquiry* 35 (April): 244–60.
- Bates, Robert, et al. 1998. *Analytic Narratives*. Cambridge: Cambridge University Press.
- Becker, Gary S. 1976. *The Economic Approach to Human Behavior*. Chicago: University of Chicago Press
- . 1996. *Accounting for Tastes*. Cambridge, Mass.: Harvard University Press.
- Bendix, Reinhard. 1960. *Max Weber: An Intellectual Portrait*. Berkeley: University of California Press
- Blute, Marion. 1997. "History versus Science: The Evolutionary Solution." *Canadian Journal of Sociology* 22:345–64
- Boswell, Terry, and William Dixon. 1993. "Marx's Theory of Rebellion: A Cross-National Analysis of Class Exploitation, Economic Development, and Violent Revolt." *American Sociological Review* 58 (5): 681–702.
- Brinton, Mary. 1988. "The Social-Institutional Bases of Gender Stratification: Japan as an Illustrative Case." *American Journal of Sociology* 94 (2): 300–334
- Brown, Cliff, and Terry Boswell. 1995. "Strikebreaking or Solidarity in the Great Steel Strike of 1919: A Split Labor Market, Game-Theoretic, and QCA Analysis." *American Journal of Sociology* 100 (6): 1479–1519.
- Brustein, William. 1996. *The Logic of Evil*. New Haven, Conn.: Yale University Press.
- Bueno de Mesquita, Bruce. 1996. "Counterfactuals in International Affairs: Some Insights from Game Theory." Pp. 211–29 in *Counterfactuals in International Relations*, edited by Philip Tetlock and Aaron Belkin. Princeton, N.J.: Princeton University Press.
- Bunge, Mario. 1996. *Finding Philosophy in Social Science*. New Haven, Conn.: Yale University Press.
- Burawoy, Michael. 1989. "Two Methods in Search of Science: Skocpol vs. Trotsky." *Theory and Society* 18:759–805.
- Burt, Ronald. 1992. *Structural Holes: The Social Structure of Competition*. Cambridge, Mass.: Harvard University Press.

- Carruthers, Bruce 1996 *City of Capital: Politics and Markets in the English Financial Revolution*. Princeton, N.J.: Princeton University Press.
- Clemens, Elisabeth. 1997. *The People's Lobby: Organisational Innovation and the Rise of Interest Group Politics in the United States, 1890-1925*. Chicago: University of Chicago Press.
- Cohen, Bernard. 1989 *Developing Sociological Knowledge*. Chicago. Nelson-Hall.
- Coleman, James. 1986 "Social Theory, Social Research, and a Theory of Action." *American Journal of Sociology* 91 (6): 1309-35.
- . 1990. *Foundations of Social Theory*. Cambridge, Mass.: Belknap Press.
- Collins, Randall. 1986 *Weberian Sociological Theory*. Cambridge: Cambridge University Press.
- David, Paul. 1985. "Clio and the Economics of QWERTY." *American Economic Review* 75:332-37.
- Dawkins, Richard. 1997. "Human Chauvinism." *Evolution* 51:1015-20.
- Dennett, Daniel C. 1995. *Darwin's Dangerous Idea*. New York: Touchstone Books.
- Elster, Jon. 1983 *Explaining Technical Change*. Cambridge: Cambridge University Press.
- . 1985. *Making Sense of Marx*. Cambridge: Cambridge University Press.
- Erikson, Kai. 1966 *Wayward Puritans*. New York: Wiley.
- Ermakoff, Ivan. 1997 "Prelates and Princes: Aristocratic Marriages, Canon Law Prohibitions, and Shifts in Norms and Patterns of Domination in the Central Middle Ages." *American Sociological Review* 62 (3): 405-22.
- Frank, Robert. 1988. *Passions within Reason*. New York: Norton.
- Friedman, Debra. 1995 *Toward a Structure of Indifference: The Social Origins of Maternal Custody*. New York: Aldine de Gruyter.
- Friedman, Debra, and Carol Diem. 1993. "Feminism and the Pro(Rational)Choice Movement. Rational Choice Theory, Feminist Critiques, and Gender Inequality." Pp. 91-114 in *Theory on Gender/Feminism on Theory*, edited by Paula England. New York: Aldine de Gruyter.
- Friedman, Debra, and Michael Hechter. 1988 "The Contribution of Rational Choice Theory to Macrosociological Research." *Sociological Theory* 6:201-18.
- Friedman, Debra, Michael Hechter, and Satoshi Kanazawa. 1994. "A Theory of the Value of Children." *Demography* 31 (3). 375-402.
- Friedman, Milton. 1953. *The Methodology of Positive Economics*. Chicago: University of Chicago Press.
- Gambetta, Diego. 1993. *The Sicilian Mafia*. Cambridge, Mass.: Harvard University Press.
- Goldthorpe, John. 1996. "The Quantitative Analysis of Large-Scale Data Sets and Rational Action Theory: For a Sociological Alliance." *European Sociological Review* 12:109-26.
- Gould, Stephen Jay. 1997. "Self-Help for a Hedgehog Stuck on a Molehill." *Evolution* 51:1020-23.
- Griffin, Larry. 1993. "Narrative, Event-Structure Analysis, and Causal Interpretation in Historical Sociology." *American Journal of Sociology* 98 (5): 1094-133.
- Hall, John R. 1988. "Social Organization and Pathways of Commitment: Types of Communal Groups, Rational Choice Theory, and the Kanter Thesis." *American Sociological Review* 53:679-92.
- Hechter, Michael. 1983. *The Microfoundations of Macrosociology*. Philadelphia: Temple University Press.
- . 1987. *Principles of Group Solidarity*. Berkeley and Los Angeles: University of California Press.
- . 1992a. "The Insufficiency of Game Theory for the Resolution of Real-World Collective Action Problems." *Rationality and Society* 4.33-40.

- . 1992b. "Should Values Be Written Out of the Social Scientist's Lexicon?" *Sociological Theory* 10 (2): 215–31.
- . 1994. "The Role of Values in Rational Choice Theory." *Rationality and Society* 6 (3): 318–33.
- Hechter, Michael, and Satoshi Kanazawa. 1997. "Sociological Rational Choice Theory." *Annual Review of Sociology* 23:191–214.
- Hechter, Michael, Lynn Nadel, and Richard E. Michod, eds. 1993. *The Origin of Values*. New York: Aldine de Gruyter.
- Heckathorn, Douglas. 1988. "Collective Sanctions and the Creation of Prisoner's Dilemma Norms." *American Journal of Sociology* 94:535–62.
- . 1990. "Collective Sanctions and Compliance Norms: A Formal Theory of Group-Mediated Social Control." *American Sociological Review* 55:366–84.
- Hedström, Peter. 1994. "Contagious Collectives: On the Spatial Diffusion of Swedish Trade Unions, 1890–1940." *American Journal of Sociology* 99 1157–79.
- Hedström, Peter, and Richard Swedberg. 1996. "Rational Choice, Empirical Research, and the Sociological Tradition." *European Sociological Review* 12 (2): 127–46.
- Hempel, Carl (1942) 1959. "The Function of General Laws in History." Pp. 344–56 in *Theories of History*, edited by P. Gardiner. New York: Free Press.
- Jasso, Guillermina. 1988. "Principles of Theoretical Analysis." *Sociological Theory* 6: 1–20.
- Kernan, Alvin. 1997. *What's Happened to the Humanities?* Princeton, N.J.: Princeton University Press.
- Kiser, Edgar. 1994. "Markets and Hierarchies in Early Modern Fiscal Systems: A Principal-Agent Analysis." *Politics and Society* 22 (3): 285–316.
- . 1996. "The Revival of Narrative in Historical Sociology: What Rational Choice Theory Can Contribute." *Politics and Society* 24 (3): 249–71.
- Kiser, Edgar, and Kathryn Baker. 1994. "Could Privatization Increase the Efficiency of Tax Collection in Less Developed Countries?" *Policy Studies Journal* 22 (3): 489–500.
- Kiser, Edgar, and Michael Hechter. 1991. "The Role of General Theory in Comparative-Historical Sociology." *American Journal of Sociology* 97:1–30.
- Kiser, Edgar, and Joachim Schneider. 1994. "Bureaucracy and Efficiency: An Analysis of Taxation in Early Modern Prussia." *American Sociological Review* 59 (2) 187–204.
- Koertge, Noretta, ed. 1997. *A House Built on Sand: Flaws in the Cultural Studies Account of Science*. Oxford: Oxford University Press.
- Kreps, David. 1990. *Game Theory and Economic Modeling*. Oxford: Clarendon Press.
- Kuhn, Thomas. (1962) 1970 *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.
- Lakatos, Imre. 1978. *The Methodology of Scientific Research Programs*. Cambridge: Cambridge University Press.
- Laudan, Larry. 1990. *Science and Relativism: Some Key Controversies in the Philosophy of Science*. Chicago: University of Chicago Press.
- Levi, Margaret. 1988. *Of Rule and Revenue*. Berkeley and Los Angeles: University of California Press.
- Liebowitz, S. J., and Stephen E. Margolis. 1990. "The Fable of the Keys." *Journal of Law and Economics* 33:1–25.
- . 1995. "Path Dependence, Lock-In and History." *Journal of Law, Economics and Organization* 11:205–26.
- Lindenberg, Seigwart. 1989. "Social Production Functions, Deficits, and Revolutions: Prerevolutionary France and Russia." *Rationality and Society* (July), 50–76.
- Lynch, Michael, and David Bogen. 1997. "Sociology's Asociological 'Core': An Examination of Textbook Sociology in Light of the Sociology of Scientific Knowledge." *American Sociological Review* 62:481–93.

- Mann, Michael. 1984. "The Autonomous Power of the State: Its Origins, Mechanisms, and Results." *Archives Européennes de Sociologie* 25:185-213.
- . 1986. *The Sources of Social Power*. Cambridge: Cambridge University Press.
- McCloskey, Donald N. 1990. *If You're So Smart: The Narrative of Economic Expertise*. Chicago: University of Chicago Press.
- Mueller, David. 1983. *Public Choice II*. Cambridge: Cambridge University Press.
- North, Douglas. 1991. *Institutions, Institutional Change, and Economic Performance*. Cambridge: Cambridge University Press.
- Parsons, Talcott. 1937. *The Structure of Social Action*. New York: McGraw-Hill.
- Popper, Karl R. 1994. *The Myth of the Framework: In Defence of Science and Rationality*, edited by M. A. Notturmo. London: Routledge.
- Przeworski, Adam. 1991. *Democracy and the Market*. Cambridge: Cambridge University Press.
- Quadagno, Jill, and Stan Knapp. 1992. "Have Historical Sociologists Forsaken Theory? Thoughts on the History/Theory Relationship." *Sociological Methods and Research* 20 (4): 481-507.
- Roemer, John, ed. 1986. *Analytical Marxism*. Cambridge: Cambridge University Press.
- Rosenberg, Hans. 1958. *Bureaucracy, Aristocracy, and Autocracy: The Prussian Experience, 1660-1815*. Cambridge, Mass.: Harvard University Press.
- Sayer, Andrew. 1992. *Method in Social Science*. London: Routledge.
- Schelling, Thomas. 1960. *Micromotives and Macrobehavior*. New York: Norton.
- Searle, John R. 1993. "Rationality and Realism: What Is at Stake?" *Daedalus* 122 (4): 55-83.
- . 1995. *The Construction of Social Reality*. New York: Free Press.
- Sewell, William H., Jr. 1996. "Three Temporalities: Toward an Eventful Sociology." Pp. 245-80 in *The Historic Turn in the Human Sciences*, edited by Terrence McDonald. Ann Arbor: University of Michigan Press.
- Skocpol, Theda. 1979. *States and Social Revolutions*. Cambridge: Cambridge University Press.
- . 1994. "Reflections on Recent Scholarship on Social Revolutions and How to Study Them." Pp. 301-43 in *Social Revolutions in the Modern World*, edited by Theda Skocpol. Cambridge: Cambridge University Press.
- Sober, Elliott. 1993. *Philosophy of Biology*. Boulder, Colo.: Westview Press.
- Stark, Rodney. 1997. *The Rise of Christianity*. San Francisco: HarperCollins.
- Stinchcombe, Arthur. 1968. *Constructing Social Theories*. New York: Harcourt, Brace & World.
- . 1978. *Theoretical Methods in Social History*. New York: Academic Press.
- . 1982. "Review Essay: The Growth of the World-System." *American Journal of Sociology* 87 (6): 1389-95.
- . 1991. "The Conditions of Fruitfulness of Theorizing about Mechanisms in Social Science." *Philosophy of the Social Sciences* 21 (3): 367-88.
- Stryker, Sheldon. 1959. "Symbolic Interaction as an Approach to Family Research." *Marriage and Family Living* 21:111-19.
- Swanson, Guy. 1960. *The Birth of the Gods*. Ann Arbor: University of Michigan Press.
- Taylor, Michael. 1988. "Rationality and Revolutionary Collective Action." Pp. 63-97 in *Rationality and Revolution*, edited by Michael Taylor. Cambridge: Cambridge University Press.
- Tilly, Charles. 1964. *The Vendee*. Cambridge, Mass.: Harvard University Press.
- , ed. 1975. *The Formation of National States in Western Europe*. Princeton, N.J.: Princeton University Press.
- Wallerstein, Immanuel. 1974. *The Modern World-System I*. New York: Academic Press.

- Weber, Max. (1922) 1968. *Economy and Society*. Berkeley: University of California Press.
- Weingast, Barry. 1996. "Off-the-Path Behavior: A Game-Theoretic Approach to Counterfactuals and Its Implications for Political and Historical Analysis." In *Counterfactuals in International Relations*, edited by Philip Tetlock and Aaron Belkin. Princeton, N.J.: Princeton University Press.
- Weingast, Barry, and Michael Moran. 1983. "Bureaucratic Discretion or Congressional Control? Regulatory Policy Making by the Federal Trade Commission." *Journal of Political Economy* 91:765-800.
- Williamson, Oliver. 1993. "Transaction Cost Economics and Organization Theory." *Industrial and Corporate Change* 2 (2): 130-46.

Limitations of Rational Choice Theory¹

Raymond Boudon
CNRS

THE ATTRACTIVENESS OF RATIONAL CHOICE THEORY

The appeal of rational choice theory (RCT) has been well explained by James Coleman (1986): rational choice theory has "a unique attractiveness" as a basis for theory because it is such a complete conception of action "that we need ask no more questions about it." Hollis (1977) has expressed the same idea in other words: "rational action is its own explanation" (quoted by Goldthorpe 1996). It is true that, once we have explained that subject X has done Y rather than Y' because Y was more advantageous, we need to know nothing more. Even if biology was able to describe adequately the chemical or electrical processes going on in the brain when a subject makes a decision, this would add nothing to the explanation as to why the subject did Y. It would merely describe the same process in a different language. But the biological explanation would be unable to confirm or disconfirm the rational explanation. This "final" aspect of rational explanations, the fact that these explanations are without "black box" frustrations, is probably, as suggested by Coleman, the main source of RCT's attractiveness.

IS RCT GENERAL?

Two questions should be raised before we give RCT the status of a general theory. Being attractive does not necessarily imply that a theory is acceptable, valid, or true in all circumstances. That "rational action is its own explanation" is one thing. Whether action can always be considered rational in the very special sense RCT gives to the notion of rationality is another. If all actions and, further, all social phenomena could be validly explained by RCT, this would be fine. But can they? In the same way, it would be just fine if I had lost my key bundle under the streetlight where I may search more effectively than in the surrounding darkness. But a preference for looking in an illuminated area does not mean that I will

¹ Address correspondence to Raymond Boudon, CNRS, 54 Boulevard Raspail, 75270 Paris, Cedex 06, France

drop my keys under the streetlight instead of in a less accessible place. So, an important question to ask is whether RCT is effectively, as contended by its followers, general. If the answer is yes, the discussion can stop at this point. If no, the next question is whether RCT can be revised to make it more general.

Can we legitimately apply RCT to all research situations and to all problems? The answer to this latter question is no for a simple reason: RCT assumes that individual action is *instrumental*, namely that it has to be explained by the actors' will to reach certain goals. Now, action can be noninstrumental, as most sociologists have recognized. Schütz, through his distinction between *Weil* and *Wozu motive*, and Weber, through his distinction between instrumental and axiological rationality, have stressed that action is not always instrumental. If the instrumentality of action is indeed limited, then RCT cannot claim to be a general theory of action. Thus, one cannot apply RCT notably in the cases where an actor does X because he believes in Z and that Z implies his doing X independently of the consequences of X. And even when action is instrumental, it can mobilize beliefs that need to be explained and that normally will not be explained by RCT. Thus, the French authorities were, for many years, more reluctant than the Dutch to use methadone to curb drug addiction. Why? Because the Dutch thought methadone was an adequate means while the French thought it was not. But *why* had the French and the Dutch different beliefs on this point?¹ RCT is of little help here.

Of course, these objections are not new and there are two traditional ways to overcome them. First, one may promote the generality of RCT by supposing that actions that appear to be noninstrumental are actually instrumental *at a deeper level*. This conversion from noninstrumental to instrumental is obtained by introducing the postulate that, contrary to appearances, beliefs are the product of self-interest. This assumption constitutes the core of some classical theories besides RCT. It was introduced notably by Nietzsche, by Pareto, and by Marx: I believe X, because believing X serves my psychological interests or my class interests.

A second way of salvaging the generality of RCT is to appeal to Milton Friedman's epistemology and to treat the causes of behavior, by principle, as unknowable. In this view, one set of assumptions about the causes of behavior is as good as any other, and it can legitimately be assumed that self-interest explains any behavior. By virtue of this "positivistic" epistemology, this assumption cannot be discussed: the only thing that matters

¹ The answer is given by Bergeron (1999) who, rather than RCT, uses the theory of rationality I advocate here.

is whether or not theories incorporating this postulate reproduce correctly the observed data.

Assumptions of type 1 can be acceptable in some cases. Thus, I can believe that social inequalities are unfair because such a belief makes my poverty undeserved and thus more acceptable to me; alternatively, I believe that social inequalities are fair because I would then perceive my opulence as well deserved. It is hard, though, to accept the view that *all* beliefs are generated by their psychological or social function. This objection was explicitly raised by Weber ([1920] 1986, p. 241) against Nietzsche and Marx: "ressentiment" theory applies exclusively to particular cases. In the general case, psychological or social interests may draw my attention to a theory and eventually create in my mind a positive or negative disposition toward the theory in question. But interests alone are generally unable to explain conviction.

Assumption of type 2 rests upon a very debatable epistemology. Why such a view as Friedman's "positivism" was developed can be understood: it derives from the reluctance of the positivistic tradition toward taking into account subjective factors. But such an epistemology is ungrounded: I can check whether this man whom I see cutting wood in his yard wants his room to get warmer. If he puts the piece of wood in his chimney, my interpretation of his behavior will be confirmed. If the weather is hot or if he starts carving the piece of wood, my interpretation will be falsified. Even if I cannot perceive directly his reasons, I can reconstruct them. This reconstruction has the status of a theory that can be confronted with data. That the reasons motivating people are not directly observable does not imply that their reconstruction is doomed to be arbitrary. Now, Kiser and Hecter reject both Friedmanian positivism and also the view that determining the reasons explaining actions would be an empirical question. Instead, they see self-interest as the ultimate *real* cause of any action. But this raises a difficult question. The rational choice theorist who encounters a voter who tells him that he votes because he considers voting to be a civic duty will reject the interpretation of the subject himself and, by application of his own RCT, assume that voting maximizes for the voter some costs-benefits balance. At this point, the rational choice theorist should explain the "false consciousness" he attributes to the actor: Why does the subject think he votes for one reason while he *really* votes for another? But how does the rational choice theorist know that the consciousness of his voter is "false"? I am not saying here that what the actor himself thinks and says of his own motivation is the ultimate truth. It is rather one piece of information among others. What I am saying is (1) that actor's statements about his motivation are *facts* that, as any fact, should be taken into consideration and explained, (2) that the Nietzschean,

the Marxian, or the rational choice theorist who attributes to the behavior of the observed subject causes that the latter does not endorse should explain his "false consciousness," notably if he rejects the Friedmanian epistemology.

On the whole, none of the strategies currently used to make noninstrumental actions instrumental appears very convincing. Each raises more questions than it answers. In other words, noninstrumental actions cannot easily be converted into instrumental actions. Consequently, RCT cannot be held as general.

Beside this basic objection, "empirical" objections can be raised against RCT, namely that it has never succeeded in explaining satisfactorily important classes of phenomena. Voting is the best known of these classes. People vote, though any individual vote has a practically zero probability of having an effect on the outcome of an election. In this case, the anticipated consequences of individual action cannot easily be taken for the cause of the action. An enormous literature has been devoted to this problem; it tries to reconcile RCT with the hard facts of voting (see, e.g., Overbye 1995). Many other puzzles in the same style could be evoked: Why do people appear so easily upset by political corruption and so sensitive to it? This familiar observation cannot be easily explained by RCT, since political corruption has, in the case of Western democracies at least, a negligible and invisible effect on the well-being of citizens. In other words, the rejection here is implausibly the effect of the consideration by social actors of the consequences of corruption on themselves. Allais (1953) and a number of authors after him have revealed another Achilles' heel of RCT (if I may make Achilles' heels plural) and shown that people do not behave effectively according to RCT predictions: when they have to choose between lotteries, in given experimental circumstances, they do not behave as maximizers.

Rational choice theorists have tried to meet these objections by auxiliary assumptions. For instance, nonvoting would include a high social cost because it would be disapproved of socially; cognitive biases make people overestimate the weight of a single vote; biases would have the effect that people see in an erroneous fashion the mathematical expectation of lotteries; cognitive "frames" predict that people do not see the world as it is. I cannot discuss these theories in detail here (I have done so elsewhere; see Boudon 1996, 1997a). In fact, none of the rational choice theories proposed to explain why people vote is satisfactory. As to the theories that propose to reconcile RCT with observed data by introducing the notions of "frames," "biases," and so on, they appear as empty ("frame," "bias" are mere words); moreover, they purchase the reconciliation with observation at a high price, since they lose the "unique attractiveness" of RCT. Once one has introduced the assumption that people see the world in such and

such fashion because a bias or a frame affects their perception, the next question is, namely, Where do these biases come from? As RCT has no answer to this question, the explanation is no more final; it generates, on the contrary, large black boxes.

The two categories of objections merge into a general objection. Some actions are purely instrumental. Among the purely instrumental actions, some are egoistic. Some actions are not purely instrumental in the sense that they include a cognitive dimension: the actor wants to reach a goal, G; he has the impression that M is a good way of reaching G, but the relation between M and G is not trivial. In that case, the noninstrumental cognitive dimension of action is the focus of the analysis. Some actions are not instrumental at all, as when an actor does X not because he wants to generate some outcome, but because X is a consequence of the principles he endorses. In that case, the main point in the analysis is to explain why the actor endorses the principles. Endorsing principles, endorsing a theory or a viewpoint is also an action, but of the noninstrumental type. This diversity cannot be forgotten or reduced except by two controversial strategies: considering the noninstrumental aspects of actions as uninteresting and being content with saying that the actors are subject to biases, frames, and so forth; or assuming that all actions would be *at a deeper level* of a unique type: not only instrumental, but egoistic. I agree with Somers that such an assumption has a metaphysical flavor.

ALTERNATIVE WAY

Instead of trying to salvage RCT against these objections, a more fruitful move is to question the basic postulates of RCT. Somers says rightly that the weak side of RCT is not its individualistic approach but its definition of "rationality": intentionality, self-interest, maximization. Do we need to endorse this very special view about rationality? Do we really need to accept the idea that all actions are not only consequential but egoistic? Is this version of rationality the only one representing the uniqueness of providing explanations without black boxes? The greatest of classical historical sociologists, Tocqueville and Weber, have implicitly answered this question. Yes, they said, action should be considered as meaningful; yes, the meaning of his action to the actor should be perceived as its cause; yes, in most cases, the meaning to the actor of his action resides in the reasons he perceives as strong to adopt it: in other words, they introduced the general postulate that the causes of an action reside in the *reasons* the actor has of adopting this action. And they added that, depending on the situation the actor is involved in, these reasons can take the form of cost-benefit considerations *but also other forms*. Thus, endorsing a theory is in most cases an action caused by the fact that one sees strong reasons of

endorsing it. Priestley endorsed the phlogiston theory, not only because he had a strong interest defending it, but because he was convinced it explained many facts better than did alternative theories.

These classical writers should be heard: they have solved implicitly by anticipation some of the crucial questions raised by the present discussion and indicated a path worth following today. They start from the view that (what *we* call) RCT can be useful notably in the situations where the actor is invited to apply a cost-benefit analysis by the very nature of the situation. But, in other circumstances, the notion of "rationality" must be given another content. In modern words, RCT is not a general theory because it uses a much too rigid and narrow conception of rationality. To illustrate, I will evoke briefly two examples from Tocqueville.

TWO EXAMPLES FROM TOCQUEVILLE

Example 1

Using implicitly "methodological individualism" and RCT.—At one point in his *Old Regime* Tocqueville ([1856] 1955) wonders why, at the end of the 18th century, French agriculture remains stagnant at a time when agriculture is flourishing in England. This is particularly puzzling since the physiocrats, who develop the view that modernizing agriculture is the main path to growth, are politically very influential in France at the time. Tocqueville's explanation: administrative centralization is the cause of the fact that positions of "civil servants" are more numerous and hence more easily available in France than in England. Also, the French centralization makes serving the king in France a unique source of prestige, influence, and power; consequently, other things equal, landlords are more easily incited in France than in England to leave their lands and buy a royal position. In England by contrast, being an innovative landowner not only produces local respect and prestige, it may also open the way to Westminster. This macroscopic difference between England and France, summarized by Tocqueville by his notion of "administrative centralization," explains why landlord absenteeism is much larger in France than in England. Further, landlord absenteeism is the cause of a low rate of innovation: since their interests are at the court, the landlords themselves have little motivation to innovate; as to the farmers who run the landownerships, they would have a motivation to innovate, but hardly the capacity of doing so. Finally, the low rate of innovation is responsible for the stagnation of agricultural development in France.

In this discussion, Tocqueville uses "methodological individualism" (MI). The macroscopic difference between France and England is explained as the effect of individual decisions taken by the landlords. The individual decisions are analyzed as taken, not by "angels," but by men

belonging to social contexts. The parameters characterizing the French and British contexts are themselves the products of a long history. This point gives me the opportunity of stressing that MI does not imply solipsism as soon as individual decisions are analyzed, as here by Tocqueville, as affected by the parameters characterizing the context.³ Finally, Tocqueville uses here what we call RCT: by leaving their land and serving the king, the landlords gain in influence, prestige, and so on. In England by contrast, it is a better strategy to appear locally as a modern and efficient landlord. The macroscopic statement "centralization is a cause of agricultural underdevelopment" appears as entirely acceptable, because it is supported by this individualistic analysis. Though "centralization" is a complex factor, it is identified with precise "parameters" that affect the situation of decision making of the actors, here the landlords. "Centralization" is a construct. But it is not a mere word. In summary, Tocqueville uses MI and, moreover, he uses the basic behavioral axiomatics of RCT; the individuals are analyzed as selfish, goal-oriented, and maximizers. It can be noted incidentally that Tocqueville's path has been literally followed by Root (1994) in his illuminating book on the comparative development of the modern state in Britain and France.

Example 2

Using implicitly MI, but rejecting RCT.—In other circumstances, Tocqueville uses MI but not RCT. Thus, he wonders, again in his *Old Regime*, why the cult of Reason became immensely popular in France at the end of the 18th century, but not in England. His answer is that traditional institutions, and hence "Tradition" with a capital *T*, were totally disqualified in France but not in England. Thus, the British aristocracy fulfilled important social and economic roles. Consequently, its higher status was considered by people as grounded and legitimate. In France, by contrast, the gentry had no visible social and economic function except sitting in Versailles. Those members of the gentry who were not able to buy a royal position remained on their land. Poor and bitter, they stuck ritualistically to their privileges. Their officially higher rank was perceived by the peasants as illegitimate. As it was the product of tradition, the peasants came to the idea that institutions deriving their strength from tradition were bad. So, when the *philosophes* proposed to substitute for institutions grounded on Tradition, a society grounded on what they presented as the opposite term, namely Reason with a capital *R*, they had

³ Bunge (1996) stresses rightly that MI (especially as I have applied it; see Boudon 1974) has nothing to do with atomism or solipsism and is perfectly compatible with his systemism.

immediate success. After all, the notion translated widespread feelings. Tocqueville makes clear that this success cannot be analyzed as the product of interpersonal influence, since it was immediate. So, the macroscopic phenomenon under examination, namely the fulgurant success of the idea of Reason, is analyzed by Tocqueville as the effect of the fact that individual French peasants, lawyers, and so on, accepted easily the theory that good institutions should be the effect of social engineering (in our language), be the product of Reason (in 18th-century parlance). This analysis follows MI, but not RCT. Here again, a social fact, in this case a difference between France and England, is analyzed as the product of reasons, but not of the RCT type. Individual peasants tend to endorse the political theory proposed by the philosophes because this theory appears to them as valid. Evidently, they expected returns for themselves from the application of this theory. Most of them were probably convinced that their condition would become better if the ideas of the philosophes were applied. But this does not explain why they saw "social engineering" as a good political philosophy. Tocqueville uses here implicitly a view of rationality that I have proposed as "cognitive rationality." The peasants endorse the political theory of the philosophes because they have strong cognitive reasons for seeing it as valid. This type of rationality is typically at work in the case of the scientist who chooses theory T against a theory T'. Cognitive rationality in my language overlaps with "rationality" in the sense that historians and philosophers of science use this word. Tocqueville's analysis provides a powerful hint here: that the type of rationality at work in the endorsement by scientists of an idea or theory is also at work in "ordinary knowledge."

COGNITIVE RATIONALITY

Cognitive rationality should be distinguished from instrumental rationality. First, because endorsing a theory is a noninstrumental action. Second, because the question the actor is confronted with here is not to maximize any cost-benefit balance, but to check whether, to the best of his knowledge, an idea is acceptable. Radnitzki (1987) has tried to reduce this cognitive rationality to RCT. When more and more facts appeared easily explainable by the theory that the earth was spherical, it became more and more difficult to develop alternative arguments supporting the theory that the earth was flat. Radnitzki proposes to substitute "costly" for "difficult" and makes then the point that the choice between alternative scientific theories can also be analyzed in RCT terms. But the main point is that the arguments supporting the theory that the earth is spherical appeared, after a while, much stronger than the arguments supporting the alternative theories. Therefore, little is gained by substituting "costly" for "diffi-

cult." "Cost" is namely a consequence of "difficulty." So, what needs to be explained is why a set of arguments appears as defensible or not.

AXIOLOGICAL RATIONALITY

Although I cannot go very far in a short note on this, another important point, as suggested by Weber, is "axiological rationality," which should be distinguished from "instrumental rationality." This notion has been much discussed, and what Weber meant by it is not clear. My own interpretation (see Boudon 1997a) is that he wanted to introduce the idea that in some circumstances actors do X not because they expect any desirable consequence, but because they are convinced that X is good, since it is grounded on strong reasons. Thus, Weber would probably never have considered voting as a paradox. It is a paradox as long as one assumes that rational action is always consequentialist, as RCT assumes. In that case, voting is well a paradox, since my vote has with quasi certainty no consequence on the outcome of the election. People vote, though. Why? Because they have strong reasons to believe that democracy is better than alternative regimes, they see that elections constitute a major institution of democracy, they understand the principle "one man, one vote," they see that this principle is an expression of a basic value, and so forth. In other words, they vote because one *should* vote if one believes in the value of democracy. This explanation mobilizes what, following my interpretation, Weber called "axiological rationality." I vote because I think I should vote. I think I should vote because I have strong reasons to believe in democracy. Of course, I will have strong reasons to refrain from voting if none of the candidates convinces me, if I do not know how I should vote, for example.

As I said earlier, many efforts have been made to explain voting by RCT theories. But these theories are all unconvincing and moreover assume implicitly without explaining it the existence of a "false consciousness" of the voter: this is, then, a huge black box.

CONCLUSION

Finally, we come to the idea that Tocqueville and Weber among others have sketched a model, which I have proposed to call the "cognitivist model" (CM), resting on the following postulates (see Boudon 1994, 1996a, 1997b, 1998).

1. Until the proof to the contrary is given, social actors should be considered as rational in the sense that they have strong reasons of believing what they believe, of doing what they do, and so forth.
2. In particular cases, these reasons can be realistically treated as deal-

ing with the difference between costs and benefits of alternative lines of action. In other cases, they cannot: in particular when a decision or an action rests upon normative or cognitive beliefs, the reasons will generally not belong exclusively to this type. This results from the fact that beliefs are unintentional, and that normative beliefs are not always consequentially grounded. Also, in many circumstances, a social actor can be personally distant from an issue, yet have strong feelings about it (a good example of this is the use of the death penalty in the United States). This point is crucial: it suggests that RCT is of limited use in the analysis of public opinion.

3. In some circumstances, the core of some action is constituted by "cognitive" reasons: he did X because he believed Z is likely or is true, and because he had strong reasons of believing so.

4. In some circumstances, the core of some action is constituted by axiological reasons: he did X because he believed that Z is fair, good, unfair, and so on, and had strong "nonconsequential" reasons of believing so.

It follows from these postulates that RCT is a particular case of CM. When the reasons in CM are restricted to belong to the benefits-minus-costs type, we get RCT. Reciprocally, when the restriction that reasons should belong to the benefits-minus-costs type is lifted in RCT, we get CM. Again, RCT is a powerful model; it cannot be held as a general theory.⁴

CM supposes that actions, decisions, and beliefs are meaningful to the actor in the sense that they are perceived by him as grounded on reasons. Even though he cannot be able to identify these reasons clearly, he has the intuitive impression that they are grounded on reasons.

Two important remarks can be introduced here. Although it is tautological to define "rationality" by the notion of "strong reasons," it is the only way of getting rid of the discussions as to "what rationality *really* means," where the discussants expose generally what *they* mean. As to the postulate that beliefs and actions are grounded on reasons, it is not tautological. Many traditions start, on the contrary, from the assumption that actions and beliefs are not the effect of reasons. As to finding out those reasons, reconstructing them, this can be a hard job: Why were the French landlords less innovative than their British counterparts? Why did Priestly believe in the phlogiston? Why did Englishmen of the 18th century believe that miners should be paid more than soldiers? (On this, see Smith 1976.) Why did the people of London not try to exert pressure on political power by street gatherings as frequently as did the Parisians? (See Root 1994.)

⁴I defended the two points already (Boudon [1977] 1982, chap. 7) but explored elsewhere the non-RCT dimensions of rationality (see Boudon [1994] and later publications)

Why are magical beliefs more likely to found in some societies than in others? (See Boudon 1994.) Why was methadone used much earlier in Holland than in France? (See Bergeron 1999.) All of these questions have been convincingly answered in works that use MI and an open theory of rationality rather than the special figure of rationality used by RCT. In particular it can be noted that, in his *Wealth of Nations*, Adam Smith, RCT's spiritual father, solves the above question about miners and soldiers using what I call the CM rather than RCT (Boudon 1996a, 1998).

The second remark, which I cannot develop, is that CM excluded radically solipsism: I cannot perceive as strong the reasons leading me to endorse a statement "X is good, legitimate, right, true, and so forth" without conceiving these reasons as grounded and hence as intersubjectively valid (Boudon 1995).

Finally, a crucial question is raised by Somers: What is a good theory? Sometimes, good scientific theories use mathematical language, are derived from a general theory, and so on. But such attributes are not components of a good theory generally. Otherwise, physical theories would be good, but biological theories bad, since the latter make little use of mathematics and are hardly deducted from a general theory. Celestial mechanics is not the model to be followed by all disciplines. A good scientific explanation of a phenomenon P is rather a set {S} of statements meeting three requirements: (1) that all $s \in \{S\}$ are acceptable, (2) that $\{S\} \Rightarrow P$, (3) that relevant facts are not arbitrarily ignored. They are satisfied, for instance, in Tocqueville's above examples: all statements are acceptable (simple psychological statements, empirical statements congruent with observation, no "black box" concepts in the components of the statements, etc.); moreover, $\{S\} \Rightarrow P$ (P = stagnation of French agriculture, enthusiasm of the peasants for the idea of Reason); finally the theory explains not only the observed behavior but the verbal statements of actors as they have reached us.

REFERENCES

- Allais, M. 1953. "Le comportement de l'homme rationnel devant le risque: Critique des postulats de l'école américaine." *Econometrica* 21 (4): 503–46.
- Bergeron, H. 1999. *Soigner ou prendre soin des toxicomanes: Anatomie d'une croyance collective*. Paris: Presses Universitaires de France.
- Boudon, R. 1974. *Education, Opportunity and Social Inequality*. New York: Wiley.
- . (1977) 1982. *The Unintended Consequences of Social Action*. London: Macmillan.
- . 1994. *The Art of Self-Persuasion*. London: Polity Press.
- . 1995. *Le juste et le vrai*. Paris: Fayard.
- . 1996. "The 'Rational Choice Model': A Particular Case of the 'Cognitive Model.'" *Rationality and Society* 8 (2): 123–50.
- . 1997a. "Le 'paradoxe du vote' et la théorie de la rationalité." *Revue Française de Sociologie* 38 (2): 217–27.

- . 1997b. "The Present Relevance of Max Weber's *Wertrationalität* (value rationality)." In *Methodology of the Social Sciences, Ethics, and Economics in the Newer Historical School: From Max Weber and Rickert to Sombart and Rothacker*, edited by P. Koslowski. New York: Springer.
- . 1998. "Social Mechanisms without Black Boxes." In *Social Mechanisms: An Analytical Approach to Social Theory*, edited by R. Swedberg and Peter Hedstrom. New York: Cambridge University Press.
- Bunge, M. 1996. *Finding Philosophy in Social Science*. New Haven, Conn.: Yale University Press.
- Coleman, J. 1986. *Individual Interests and Collective Action. Selected Essays*. Cambridge: Cambridge University Press.
- Goldthorpe, J. 1996. "Class Analysis and the Reorientation of Class Theory: The Case of the Persisting Differentials in Educational Attainment." Manuscript. Nuffield College, Oxford University.
- Hollis, M. 1977. *Models of Man: Philosophical Thoughts on Social Action*. Cambridge: Cambridge University Press.
- Overbye, E. 1995. "Making a Case for the Rational, Self-Regarding 'Ethical' Voter . . . and Solving the 'Paradox of Not Voting' in the Process." *European Journal of Political Research* 27 369–96.
- Radnitzky, G. 1987. "La perspective économique sur le progrès scientifique: Application en philosophie de la science de l'analyse coût-bénéfice." *Archives de Philosophie* 50 (avril-juin): 177–98.
- Root, H. L. 1994. *The Fountain of Privilege: Political Foundations of Economic Markets in Old Regime France and England*. Berkeley and Los Angeles: University of California Press.
- Smith, A. (1776) 1976. *An Inquiry into the Nature and Causes of the Wealth of Nations*. Oxford: Clarendon Press.
- Tocqueville, A., De. (1856) 1955. *The Old Regime and the French Revolution*. New York: Doubleday.
- Weber, M. (1920) 1986. *Gesammelte Aufsätze zur Religionssoziologie*, vol. 1. Munich: Mohr.

Initial Conditions, General Laws, Path Dependence, and Explanation in Historical Sociology¹

Jack A. Goldstone
University of California, Davis

Much of the training of the modern economist tends to weaken the trainee's natural, intuitive understanding of historical causation, so that some remedial work is required in addressing an audience, some of whose members' advanced education will have left them incapacitated in this particular way. (Paul A. David)

Margaret Somers, Edgar Kiser, and Michael Hechter are serious heavyweights in the debate on methods of explanation in historical sociology, with major publications prior to the current symposium (Skocpol and Somers 1980; Somers 1996; Kiser and Hechter 1991; Hechter 1992; Kiser 1996). But in this symposium, they have come out swinging wildly. Were I to referee this bout, I would have to penalize both parties for low blows. Somers criticizes an older and cruder form of rational choice theory (RCT) than most practitioners currently use; Kiser and Hechter in reply propose some rather odd criteria for selection among different explanations. In addition, both parties are guilty of some rather severe misunderstandings of science, and of principles of scientific explanation. We shall sort through these errors below. After these errors are cleared away, I hope to lay out how, far from there being one right way to approach explanation in historical sociology, scholars seeking to study historical phenomena need to be aware that *different forms of explanatory principles, differently emphasizing the role of initial conditions, general laws, and path dependency*, are necessary to explain different kinds of historical relationships.

A FALSE DICHOTOMY

Somers begins by laying out two "Kuhnian trajectories." One, which she claims is followed by Kiser and Hechter, is "theory centric" and insistent

¹ Direct correspondence to Jack Goldstone, Department of Sociology, University of California, Davis, California, 95616.

on deductivist general theorizing. The other, which Somers claims to draw out herself and follow, is based on "historical epistemology" and insists on problem-driven, relational and historical theorizing. This is a false dichotomy on two counts: there is no basis for this division in Kuhn's (1970) argument, and both the "theory-centric" and "historical epistemology" modes of explanation are straw men, incapable of standing in for sensible scientific historical explanation.

Somers suggests that theory-centric views of the scientific method are dominant in textbooks. Noting that "few experimenters in physics [are] given the Nobel Prize," and taking literally Kuhn's pronouncement that theories are only overthrown by other theories, not by experimental observations, Somers believes that experiment and observation have somehow been downgraded to secondary status in the arena of scientific knowledge and that "theories" are the dominant element of scientific discourse and explanation. This is simply and terribly mistaken.

First, the vast majority of Nobel Prizes in physics are and have been given for experimental discoveries or for the perfection of new experimental apparatus. From the first prize given in 1901 to Wilhelm Roentgen for the discovery of x-rays, to the 1997 prize given to Steven Chu, Claude Cohen-Tannoudji, and William Phillips for the development of methods of supercooling atoms with laser light, most winners have been experimenters, some detecting things (e.g., new elementary particles) that were predicted by theories, but more often finding things unanticipated by theories (although supremely relevant, of course, for existing and future theoretical work). Of the 10 most recent prizes (1988–97), eight went to *experimenters* for either new discoveries or the development of new experimental methods. Even that most famed general theorist of this century, Albert Einstein, did not receive his Nobel Prize for either the special or general theory of relativity. Rather, Einstein won for extremely *problem-driven* theoretical work, namely on the photoelectric effect.

I had the good fortune to discuss this matter with Arno Penzias a few years back.² He pointed out that the role of experiment in physics is much misunderstood, and that theoretical development has been abstracted from real scientific work, thanks in part to populist misreadings of Kuhn. In fact, experiment, not theory, dominates normal science. Moreover, there really is no such thing as "general" theory in physics. There are only degrees of generality. Kepler's laws were a completely problem-driven

² Somers (n. 8, in this issue) cites Penzias's discovery of the low-temperature background radiation that pervades the universe, which provided the first experimental evidence for the big bang theory of the origin of the universe, as a case of an experiment whose significance has been reduced in science texts to emphasize the role of theory.

theory that focused on finding regularities in the shape and velocity of planetary orbits in our solar system. Newton developed a more general theory of the relative motions of gravitationally bound bodies, but that too was driven by the problem of explaining how the motions of planets and terrestrial bodies could be explained by common principles. Newton's theories of motion, however, were found inadequate to deal with the behavior of light; that required new theories of electromagnetic radiation, which were eventually developed by Maxwell and others. But in the late 19th century, an experiment on the velocity of light waves by Michelson and Morley suggested that something was wrong with Newton's view of absolute reference frames for the measurement of motion in space. Kuhn was quite right that this experiment did not suddenly force physicists to "abandon" Newton's theories and the many explanations based on them. But the importance of the experiment was widely noted (Michelson received a Nobel Prize for his experiment in 1907) and many different theorists and experimenters sought to explain Michelson and Morley's results (Mason 1962, pp. 542–45). Almost two decades later, Einstein offered his special theory of relativity as a solution to the problem posed by the troubling experiments conducted by Michelson and others; it took another decade to work out the implications that relativity would have not only for Newton's laws of motion, but for Newton's laws of gravity. In the end, Einstein's laws of motion and gravity were more general than Newton's in that they applied to motion at extremely high speeds or in extremely high gravitational fields. But of course, even Einstein's theory is not wholly general, for it remains separate from and unreconciled with the laws of quantum mechanics.

Moreover, and this is the key point, Einstein did not develop his theory of relativity in response to a call to "generalize" Newton's theory of motion. Einstein's work—on the photoelectric effect, on Brownian motion, and on special relativity, all published in the same year, 1905—was highly problem driven. Each was aimed precisely at explaining the troubling results of well-known experiments that were inconsistent with the predictions of current theories. Kuhn did point out that the transition from one theory to another in science is not a smooth, quick process, in which scientists act as impartial judges, reviewing evidence and quickly throwing out a theory as soon as problematic or disconcerting evidence appears. Theories have partisans and past successes, and thus continue to be relied on until alternative theories are developed that can reconcile past results with the new problematic evidence. But that does *not* mean that such disconcerting evidence is swept under the rug or is marginalized in the evolution of science (remember Michelson received the Nobel Prize, not Einstein, for the work that eventually led to special relativity). Kuhn never sought to imply that experimental evidence was not of critical importance, or

that transitions between theories (or paradigms) were not intimately connected to experimental evidence, or that science is not essentially problem driven, for none of these things is remotely true.

But if scientific investigation and theoretical development are mainly problem driven, not theory centric, the label that Somers seeks to pin on Kiser and Hechter has no foundation. What, then, is the difference that lies between the Somers and the Kiser/Hechter positions?

INITIAL CONDITIONS, GENERAL LAWS, AND PATH DEPENDENCE

I believe the difference that is fueling this debate is simply a difference in the foregrounding given to *initial conditions* by Somers and to *general laws* by Kiser and Hechter. Somers seems to believe the importance of historically specific conditions (which I shall refer to as "initial conditions") is so great as to somehow invalidate the applicability of "general laws." She instead refers to "mechanisms" as the causal elements in narrative or "relational" explanations, which she identifies with path dependency and finds similar in operation to evolutionary biology, which she also claims has no "true invariant laws" (p. 771).

Kiser and Hechter, on the other hand, seem to fear that an obsession with initial conditions will lead to purely narrative explanations of particular sequences of events. Fearing that this "new type of research in comparative historical sociology" tended toward Dr. Seuss-like explanatory principles—"it just 'happened to happen,' and was not very likely to happen again" (Geisel and Geisel 1991, p. 91)—Kiser and Hechter sought to emphasize the key role of general theory in providing powerful causal understandings of diverse events. They argue that sociology needs theories that hold for a wide range of initial conditions. Suggesting that rational choice theory could serve this role, they sought to redirect the efforts of comparative and historical sociologists toward the explicit use of general laws in explaining historical events.

Somers's position has so many faults it will be difficult to correct them all, but we should try. First, mechanisms are not the functional equivalent of "general laws" and in no way substitute for them in explanation. A general law is a statement of necessary or probable connection between two kinds of events (e.g., for the law of reflection, the angle of incidence of light striking a reflecting surface is equal to the angle of exit from the surface). A general law is *not* general because it applies to a wide variety of different kinds of cases and events. It may only apply to a *specific* kind of problem. What makes a general law "general" is that it applies to a range of initial conditions and asserts a necessary or probable connection between particular initial conditions and a subsequent event or events. Thus the law of reflection is a general law because it applies to a wide

range of angles of incidence for light striking a reflective surface. If the exit angle varied randomly such that the angle of incidence had no relation to the angle of exit, there would be no law of reflection. However, discovering or positing the "mechanism" by which the interaction of light and a reflective surface produces an exiting beam of light with the identical angle to the angle of incidence is quite a different matter than stating the general law. Positivists may state the general law, which accounts for all observations, and be content at that. Realists, however, may wish to explain further what events or principles govern this result. (The realist answer, as conveyed to me by no less an authority than Richard Feynman, is that the light actually exits the surface at a variety of angles, as a quantum probability wave, but that the interaction that occurs when we observe the reflecting light forces the light's energy to collapse onto the fastest path from the point of emission to the reflecting surface to the point of observation—and this fastest, or shortest, path, is the one in which the angle of incidence equals the angle of exit). Kiser and Hechter too are guilty of confusing general laws and mechanisms as elements of explanations, but we shall deal with them below.

Second, because of this confusion, Somers's claim that "relational" accounts can avoid general laws and use mechanisms to build causal explanations is misguided. Narrative and relational accounts, however problem driven and historically specific, cannot rise above the wholly contingent and unique Seussian explanation—it just happened that this happened first, then this, then that, and is not likely to happen that way again—unless there is some assertion of a *necessary or highly probable* connection between events A, B, and C in the series of events that make up the narrative. Without the assertion of a necessary or probable connection, there is *no* causal account—it just happened that way. But how can one assert that initial condition A leads with high probability to subsequent condition B without a general law connecting type A and type B events?

So even relational accounts must rely on general laws. Why then does Somers think such laws can be avoided? My teacher George Homans, one of the greatest historical sociologists (see, e.g., Homans 1968), also wrote about explanation in sociology (Homans 1967). He argued that the "general laws" of human behavior were mainly psychological, and therefore of little interest to sociologists. He suggested we can take for granted—as rational choice theorists would have us do—that people seek to maximize their well-being, usually defined in terms of wealth, power, or status, through their interactions with other people and their environment. The problem with such general laws, however, is that they tell us very little about why particular national histories, for example, turn out the way they do. How do we go from knowing that people seek to increase their power to explaining that England's economic and military power sur-

passed that of Holland in the 18th century and even that of France in the 19th century? The answer is that we must trace the action of this particular principle through the action of *particular historical actors in particular historical settings*. We need to know of the economic and military resources available to the rulers of England, Holland, and France, the coalitions and conflicts they faced with their elites, and the political and economic support they received from their populations. These "initial conditions" are rather more complex and striking than the "general law" that people seek to increase their power and wealth. Indeed, apparently Somers finds the initial conditions so much more interesting and decisive that she falsely believes she can causally connect them to later outcomes without any general laws at all—a problem shared by many historians.

Third, Somers compounds her confusion about mechanisms and laws by asserting that the property of "path dependence" is equivalent to sensitive dependence on initial conditions. Now as Paul A. David (1997) has pointed out, this is a common error, but it wholly vitiates Somers's argument. Path dependence is a property of a system such that the outcome over a period of time is *not determined* by any particular set of initial conditions. Rather, a system that exhibits path dependency is one in which outcomes are related stochastically to initial conditions, and the particular outcome that obtains in any given "run" of the system depends on the choices or outcomes of *intermediate events* between the initial conditions and the outcome. The classic illustration of such a system is that of a Polya urn experiment (Arthur, Ermoliev, and Kaniovski 1983). To oversimplify, imagine an urn with four balls in it—one red, one yellow, one white, and one black. The object is to fill the urn by selecting one ball and then replacing it along with two more balls of whichever color is chosen. Which color will dominate the full urn's contents? Note that whichever ball is drawn first—red, yellow, white, or black—will gain an advantage in future rounds, for there will then be three of that color, and only one of each of the others. Therefore, 50% of the time, the first color chosen then will also be chosen second, thus receiving an even larger advantage. Nonetheless, this does not mean that the first color drawn will always fill the urn. There is a 50% chance that another color will be drawn on the second round, thus restoring parity between at least two colors, and leaving it to later choice to tip the balance in any one color's favor. The study of this class of problems has shown that there is no determinate outcome; rather the final pattern depends on the particular choices that happen to be made in the sequence of filling the urn.

To use Somers's own example, if "14th-century legal institutions . . . can be shown *under certain initial and subsequent conditions*" to not simply disappear "but to become causal factors in the development of 19th-century democratic institutions" (p. 768; my emphasis) what exactly is she

asserting? Is she saying that if we had run this process many times, under the specified initial and subsequent conditions, then in a preponderance of these runs the 14th-century legal institutions would "cause" 19th-century institutions to develop in a certain democratic way? If so, she is asserting a general law (a law of "historical sedimentation") to the effect that earlier institutions inevitably leave their impress on subsequent ones. To return to the principle of the urn, she is asserting that the *particular* institutions and characteristics that she highlights, among all those in the "urn" of 14th-century England, *had* to be causally connected to particular 19th-century outcomes—they could not disappear. But she is *not thereby* specifying a *mechanism* for that general law—we might still ask whether the sedimentation is unavoidable because institutions can only be changed incrementally, because humans are creatures of tradition and habit, or by what other processes do five century-old institutions inevitably leave their imprint on later ones. Somers's own example, in other words, seems to invoke general laws, does not specify mechanisms, and discusses what is clearly *not* a path dependent process, otherwise the possibility that 14th-century legal institutions might indeed have simply disappeared would have been an alternative path for the system.

Now, it could be that Somers has simply mischosen her words, and she means to say something different: "Given certain choices, actions, and events that were not necessary, but happened to have been made, 14th-century legal institutions in England had a decisive impact on the shape of 19th-century democratic institutions." If this is her intent, she would be telling a particular narrative in which a path-dependent system, which could have gone in many different directions from its 14th-century initial conditions, went in one particular way because of the impact of certain "subsequent" conditions that are really not conditions at all, but discrete acts or events. It is certainly true that, in path-dependent systems, the present state is a reflection of prior events and that knowing those events can tell us why the present state, and not other possible states, currently exists and is sustained. This is consistent with Somers's insistence on narrative as her preferred mode of analysis (Somers 1994). But this mode of explanation *cannot* do what Somers claims here, which is to "help to explain why some aspects of the social world, and not others, became modern in the first place" (p. 769). The last words are crucial. Tracing the evolution of a path-dependent system can tell us why certain phenomena and not others finally emerged. But only a *determinate* or *causal* system governed by general laws can tell us why certain phenomena and not others became possible *in the first place*. Only general laws would create a situation in which 14th-century initial conditions have to affect 19th-century outcomes. In a path-dependent system, a wide variety of outcomes are by definition possible *in the first place*. One can eliminate certain of

those outcomes at the outset by causal laws, or one can eliminate certain of those outcomes in the process of system change in response to subsequent events, but one cannot explain why anything had to happen in the first place solely by reference to the operation of a path-dependent system.

It is for this reason that the study of evolution, although certainly the study of a path-dependent process, *does* utilize invariant laws, as well as allowing for the contingent role of key but undetermined events. The critical law for the evolutionary process in Darwinian science is the law of natural selection, which posits that, other things equal, individuals with a genetic characteristic that aids in their survival will pass that characteristic to their offspring, who will appear in the next generation in greater frequency than the offspring of individuals who lack that characteristic. Spoken of in terms of greater reproductive success, or the spread of desirable traits throughout a population, this principle is sometimes not firmly described as an "invariant law." But if this principle were not true for all species, at all times, then Darwinian evolution could not govern the origin and evolution of species.

Darwinian evolution in fact has a number of laws that act precisely to rule out certain subsequent events *in the first place*. For example, because the transmission of traits for multicellular organisms is specified to be the result of reproduction (not of acquired characteristics nor of transmission between unrelated individuals) many pathways are ruled out, in a way they are *not* ruled out for the evolution of bacteria (which can directly exchange genetic materials) or culture (in which new characteristics can be acquired and transmitted to many unrelated individuals.) Darwinian selection, for example, although a path-dependent process that allows many particular outcomes, has *general laws* of the transmission and selection of traits that make it impossible for new genetic characteristics to be transmitted to human populations (with their 20–30 year generational time) as rapidly as to many insect populations (with generational time intervals measured in months or even days). These laws allow geneticists and evolutionary biologists to estimate the rate of continental drift, or to judge how long Australia has been isolated, by studying the differences in speciation and genetic information between the fauna in various world regions. Contrary to Somers's assertion that "Evolutionary biology . . . lacks invariant laws," such invariant laws are at the core of the enterprise and allow evolution to be studied in a scientific manner. To be sure, evolutionary biology lacks singular determinant outcomes. But that does not mean it lacks invariant laws, only that such laws work in combination with a variety of initial conditions and that contingent events (e.g., an extraterrestrial fireball ending the reign of the dinosaurs) may reshape those conditions and produce different specific outcomes at any time.

Where I fully agree with Somers is that evolutionary biology is a fine

model for the historical social sciences. Done well, evolutionary biology utilizes invariant laws and principles but combines them with a sensitivity to the effect of a wide range of initial conditions and the role of contingent events in altering those conditions over time (see Gould 1983, 1989).

If both of Somers's claims that (1) there is a theory-centric approach to science that downplays concrete observations and (2) she has a "relational" alternative that can find causal explanations in "mechanisms" and "path-dependent" narratives instead of general laws are false, what is there to her debate with Kiser and Hechter?

Much of her criticism seems to be directed at Kiser and Hechter's claims that rational choice theory is a good example of a general theory that can help sociologists explain historical change, and, more particularly, that Kiser and Hechter can see no other general theory as a clear competitor for RCT. Now Kiser and Hechter do also make a muddle of general laws and mechanisms, and make some other rhetorical errors in their 1991 article that are moderated or corrected in this symposium. But I believe Somers has misrepresented or misunderstood Kiser and Hechter's claims.

RATIONAL CHOICE THEORY AND HISTORICAL EXPLANATION

First of all, it should be understood that the predictions of rational choice theory are in fact quite sensitive to initial conditions. Indeed, much of the thrust of rational choice theory is finding out how, given certain initial distributions of rewards and penalties for certain actions and certain patterns of interpersonal interaction (e.g., simultaneous play, alternate play, repeating play), and given certain assumptions about the motivations for actors' behavior, the choices of actors will produce particular outcomes. Some of these outcomes may be to actors' advantage (cooperation), other outcomes may not (e.g., prisoner's dilemma). Modest changes in initial conditions of either rewards or information can lead RCT to predict quite different outcomes or characteristics of interaction (Schelling 1960). Rational choice theory certainly has just as much ability to explain the arrival at and maintenance of suboptimal states as do narrative accounts pointing to path dependency. Indeed, these modes of explanation are not at all adversarial; economists such as Paul David and W. Brian Arthur rely on rational self-interest to explain why certain equilibria arrived at through path-dependent processes, including the famous QWERTY keyboard, are sustained.

I suspect that Kiser and Hechter would in fact not quarrel in principle with Somers's argument that 14th-century legal institutions had a substantial impact on the shape of 19th-century English democracy. However, Kiser and Hechter would tell this story somewhat differently, in terms of individuals between the 14th and 19th centuries making a series

of choices, based on efforts to maximize their well-being in the face of "certain initial and subsequent conditions," that resulted in certain 19th-century democratic institutions emerging as the dominant choice of English political elites. Kiser and Hechter would thus see the 19th-century outcome as a path-dependent process that could have turned out otherwise, whose explanation requires tracing the specific choices faced by groups of individuals at specific times (exactly what Brustein [1996] does in discussing the reasons for the expansion of the Nazi party in Germany in what Kiser and Hechter point to as a model RCT historical work).

It is unfortunate that Kiser and Hechter too have confused issues of "general laws," "mechanisms," and desirable principles of explanation. For example, Kiser and Hechter asserted in their 1991 article and repeated in this symposium that "three criteria can help us assess explanations based on rival causal mechanisms: plausibility, reducing the time lag between cause and effect, and testing each mechanism's unique empirical implications" (p. 790). Except for the last, I do not know of any scientific basis for these criteria. Certainly quarks and superstrings do not pass any tests for plausibility, nor did black holes or continental drift or relativistic or quantum effects. (Neils Bohr and Albert Einstein continued to find many quantum principles wholly implausible long after they had been firmly established!) Nor is the time lag between cause and effect relevant. Let us say that Jane and Joe have a fight on Tuesday and that a bus honks loudly outside their house at 4 P.M. on Thursday; Jane grabs a knife and stabs Joe at 4:30 P.M. on that same day. Would a theory that the honking of the bus drove Jane mad strike us as a better explanation than the rival theory that her action was due to the fight on Tuesday? It does, after all, fare better on the criterion of "reducing the time lag between cause and effect." No doubt, it is desirable for a mechanism that links events to fill in the temporal gaps between cause and effect, and indeed often what we mean by a "mechanism" is precisely the pattern of events that links a distant cause to its necessary effect. (Perhaps this is why Somers confuses having a "mechanism" to link 14th- and 19th-century institutions with having an explanation, although mechanisms to link causes and effects are accompaniments to general laws, not to contingent outcomes.) But in no way should temporal proximity of events (as opposed to temporal *order*) be a criterion for choosing between rival causal theories.

While much of what Kiser and Hechter claim is unexceptional, it is this sort of confusion that helps fuel debates. Similarly, Somers may be confusing the claim of RCT that individuals seek to maximize their personal outcomes with the claim that social groups seek to maximize their outcomes and do so regardless of their history. If this were so, then there might be a macroclaim on the part of such a theory that 19th-century English political institutions were chosen by the English elites because

they were the most efficient for maximizing their well-being and that prior elements and issues in English history had little relevance. Now, if Somers were arguing the different macrocausal theory that 19th-century English political institutions *had* to have certain features because of the presence of certain features in England's 14th-century legal institutions, then we would indeed have two sharply opposed macrotheories of 19th-century English democracy. One theory would say those institutions were chosen for maximal efficiency in attaining elite goals, not because of any historical events that mattered; the other (Somers's?) would say those institutions were tied to centuries-earlier institutions by some historical causal processes that made the linkage unavoidable.

I have some problems asserting that this is Somers's theory, since she seems so concerned to avoid explanations based on general laws and to favor path-dependent models in which the linkage between 14th- and 19th-century institutions *would* be avoidable. But if Somers does believe that laws of "sedimentation" and incremental institutional change (similar to the laws of evolutionary biology) do hold in social history, then perhaps she believes that there is a *necessary* and not merely contingent connection between the 14th- and 19th-century institutions in England and, if so, that would account for her dispute with a theory that could be interpreted as saying that only current efficiency of outcomes, and not historical processes, matter.

If that misunderstanding is the source of the dispute, perhaps we can lay it to rest. Rational choice theory is sensitive to initial conditions and path-dependent processes and is not in conflict with giving these factors due attention. Even asserting the maximization of self-interest as an invariant principle of human behavior (although it may strike some as churlish) is fully consistent with allowing historical and contingent changes in initial conditions to determine what the eventual social outcomes of that behavior will be—for good or ill.

But I fear that is not the sole source of the problem. Rather, what Somers and other critics of RCT are reacting to is the tendency of some practitioners of RCT to grossly simplify the actual complexity of initial conditions in order to make deterministic calculations of social outcomes. For in fact, RCT's ability to reach firm conclusions from given initial conditions and patterns of interaction is limited. Many sets of initial conditions and interactions are indeterminate—like the path dependence of the Polya urn. For most RCT theorists, such problems are uninteresting—rather than tracing the steps that led to a particular outcome that may only happen once, many RCT theorists would prefer to find the solution to a class of problems, and finding a determinate solution may require severe and even unrealistic restrictions on the initial conditions and interactions.

This difference in interests is precisely what causes the conflict or mis-

communication between theoretical economists and economic historians remarked upon by Paul David (1997) in the quote that opens this essay. All modern economists, whether they are neoclassical theorists or economic historians, are firmly in the RCT camp. But the theorists seek to solve problems with determinate equilibria solutions; the economic historians are interested in particular outcomes (e.g., the Industrial Revolution in England or the standardization of the QWERTY keyboard) that may have been flukes of particular choices. Thus the theorists concentrate on solving complex problems whose initial specifications are sometimes found to be unrealistic and irrelevant by economic historians, while the economic historians concentrate on explaining particular outcomes that the theorists find to be of no great interest. Thus, to point to the most egregious example, until just over a decade ago, economic historians' major preoccupation was explaining economic growth, while economic theorists put that issue aside as depending on conditions (such as increasing returns) that seemed, at the time, to void an equilibrium solution (Arthur 1994).

A similar difference in vantage points may be separating RCT theorists operating in historical sociology from Somers and other critics of rational choice. One example from my own field, the study of revolutions, is the response to Gordon Tullock's (1971) work on the "paradox" of revolutions. Tullock showed that if individuals made separate choices as to whether to join a revolutionary movement, *and* if the benefits of a revolution went to everyone and not just those who rioted or rebelled, *then* revolution would be impossible, as a rational calculation of self-interest would always lead individuals to try to free ride on the efforts of others to create a revolution rather than to incur the risks and costs of revolutionary activity themselves. The reasoning was that for any one individual, his or her action would make no difference to the success of the revolution. Since this individual would enjoy the benefits of a revolution whether or not he or she participated, but would incur risks and costs only through participation, then he or she should certainly *not* participate. And if every individual calculated in this fashion, revolutionary mobilization could not occur.

The reaction to this result was twofold. Economists championed the result as a triumph of logic and a fascinating result. Historians and sociologists who actually studied revolution ignored it, since they knew full well that revolutions do occur, and their goal was to explain this social fact, not to posit conditions under which it would not be possible.

However, this bifurcation was not necessary and now is starting to be overcome by sociologists who have shown that the premises regarding the initial conditions, not the reasoning, in Tullock's analysis were faulty (see, e.g., Opp, Voss, and Gern 1995; Goldstone 1994). Individuals do not de-

cide to join, or not to join, revolutionary movements as isolated individuals. Instead, they decide within the dynamics of groups to which they have prior commitments and within which it appears likely that their actions will make a difference. Once the role of groups is added to the rational choice analysis, such analysis, leads to results—such as the necessity of cross-class coalitions, the role of neighborhood ties, and the development of two different pathways to revolution (one based on urban uprisings in the capital, the other based on guerilla mobilization in the countryside) corresponding to different kinds of regimes—that confirm and provide microlevel foundations for the results of macrosocial studies of revolution (Goldstone 1994). For their part, rational choice theorists are exploring themes such as the institutional and historical ways to avoid the free-rider problem (Lichbach 1995), the bases of group action and group solidarity (Hechter 1987), and the real-world implications that arise from the frequent concealment or falsification of preferences (Kuran 1995).

In sum, Tullock worked with one set of initial conditions that yielded a firmly determined, if negative and paradoxical, solution. But in the real world, the initial conditions that may give rise to revolution or revolt are more complex and involve a good bit of uncertainty among the actors (Goldstone 1989, 1994). Many different sets of initial conditions may give rise to rebellion or revolt, depending on the actions of rulers and rebels; but as Lichbach (1995) shows, these conditions are harder to specify than those chosen by Tullock. Further, once specified, conditions neither form a simple closed set, nor do they produce a simple determinate solution. In fact, discovering precisely which initial conditions give rise to which kind of rebellion or revolution, and with what outcomes, has been a central issue for the comparative historical study of revolution (see Moore 1966; Tilly 1978; Eisenstadt 1978; Skocpol 1979; and Goldstone 1991).

PRINCIPLES OF HISTORICAL EXPLANATION

If Kiser and Hechter and Somers are not truly in conflict over two different methods of historical explanation, does that mean there is one best mode that everyone can agree on? That would be, and is, too good to be true. Historical explanation is complex, and an enormous literature has been devoted to it over many decades (for a recent survey, see Lloyd [1993]). But one point I would like to make is that *good* historical analysis is distinguished not by any one method, but by choosing the method of explanation best suited for its explanatory goal.

If one is concerned to explain a historical phenomenon that emerges quite commonly, yet from a variety of different settings or initial conditions, then it would be bad practice to use a path-dependent model. What we appear to have in such cases is convergence on an equilibrium solution

from a variety of origin points. In these cases, looking for a solution based upon rational choice theory—or any other theory that leads to selecting a particular outcome as optimal from a broad set of possible outcomes and starting points—would be wise. For example, Kiser (1994) finds that, despite widely differing histories of state formation, nearly all early modern states used tax farming to collect central government revenues. It is this kind of historical situation that is most plausibly explained by solution to a common optimization problem—that of maximizing revenues under conditions of expensive monitoring and scarce information.

However, if one is concerned to explain a historical phenomenon that emerges only occasionally, but from recognizably similar initial conditions, then one should seek a general law connecting particular initial conditions with particular outcomes. Thus, for example, Reuschemeyer, Stephens, and Stephens (1992) found that mass democracy has emerged in some places but not others and only where workers or small farmers formed strong political parties that gained support from bourgeois elites. They thus argue that they have found a general principle (they shrink from calling it a law, as too deterministic sounding for historical studies) that strong parties that create cross-class coalitions are a crucial condition for the emergence of mass democracy. Skocpol (1979) found that social revolutions emerged in some places and not in others and traced this to the general principle that social revolutions in agrarian-bureaucratic states occur only where one finds the combination of international competitive pressure on the state, leverage from an autonomous elite, and organization among the peasantry.

Hechter and Kiser (1991) challenged Skocpol's analysis as *not* being an approach based on general theory, but they are wrong. By asserting a necessary connection between a set of initial conditions and a particular outcome, her argument is precisely one of developing a general law or principle that shapes specific patterns in history. The *scope* of the theory may be narrow—it fails to hold for Iran, for example, and other cases where urban groups were more important than peasants in revolutionary activity (Farhi 1990). But as a *form of explanation* it rests as much on the principles of general laws as Kiser's explanation of tax farming. It simply takes on a different type of problem than Kiser did, one in which the outcome is somewhat rare and found only with a particular set of initial conditions, and thus aspects of the explanation must perforce appear different. Kiser and Hechter may rightly claim that the scope of RCT—which is held to obtain for all individuals at all times—is much greater than the scope of Skocpol's theory, which only obtains for pre-industrial agrarian-bureaucratic states. And unlike the other criteria that they adduced, the scope of a theory (what is sometimes popularly called its "gener-

ality") is a widely used criteria for choosing among rival theories. But they cannot claim that explanations based on RCT are based on a general theory and historical macroexplanations like Skocpol's are not; that is simply not a valid basis of distinction.

Finally, if one is concerned to explain a particular unique event, one that has occurred only once and then perhaps diffused or spread but did not repeat, despite similar initial conditions being found elsewhere, then one has most likely identified a path-dependent system in which the unique outcome was produced by some contingent conditions or choices that separated the outcome in that particular system from outcomes in other systems that started from similar conditions. One example of such an event is the discovery and implementation of steam power for mechanical energy in 18th-century Britain. Despite advanced organic economies having existed for many centuries with similar resources of wood, coal, and water, and similar histories of economic growth and technological discovery, only Britain independently developed steam power (Wrigley 1988). This suggests that the explanation for this event, and for the Industrial Revolution that steam helped power, lies in some contingent event that threw the path of technological evolution onto a different track in Britain than was the case elsewhere in the world (Mokyr 1993; Goldstone 1987, 1998), and not in the solution of a common optimization problem nor in a necessary outcome of initial conditions.

Historical analysis based on RCT, on macrolevel general principles, and on identification of unique path-dependent contingencies are thus all possible, valid modes of historical explanation. What is important, then, is to avoid conflating these modes or claiming that one is "best" to the exclusion of others. What matters for *good* historical explanation is that whichever mode of explanation is used, it is wisely applied with due regard for the different roles of general laws and of varying initial conditions and of path-dependent processes in shaping various patterns of historical events. If the papers in this symposium make that clear, they will have done historical sociology a great service.

REFERENCES

- Arthur, W. Brian. 1994. *Increasing Returns and Path Dependence in the Economy*. Ann Arbor: University of Michigan Press.
- Arthur, W. B., Y. M. Ermoliev, and Y. M. Kaniovski. 1983. "A Generalized Urn Problem and Its Applications." *Cybernetics* 19:61-71.
- Brustein, William. 1996. *The Logic of Evil*. New Haven, Conn.: Yale University Press.
- David, Paul A. 1997. "Path Dependence and the Quest for Historical Economics: One More Chorus on the Ballad of QWERTY." Manuscript. All Soul's College, Oxford University

- Eisenstadt, S. N. 1978. *Revolution and the Transformation of Societies*. New York: Free Press.
- Farhi, Farideh. 1990. *States and Urban-Based Revolutions: Iran and Nicaragua*. Urbana: University of Illinois Press.
- Geisel, Theodore S., and Audrey S. Geisel. 1991. *Six by Soss*. New York: Random House.
- Goldstone, Jack A. 1987. "Cultural Orthodoxy, Risk, and Innovation: The Divergence of East and West in the Early Modern World." *Sociological Theory* 5:119-35.
- . 1989. "Deterrence in Rebellions and Revolutions." Pp. 222-50 in *Perspectives on Deterrence*, edited by P. Stern, R. Axelrod, R. Jervis, and R. Radner. New York: Oxford University Press.
- . 1991. *Revolution and Rebellion in the Early Modern World*. Berkeley and Los Angeles: University of California Press.
- . 1994. "Is Revolution Individually Rational?" *Rationality and Society* 6:139-66.
- . 1998. "The Problem of the 'Early Modern' World." *Journal of Economic and Social History of the Orient* 41:249-84.
- Gould, Stephen Jay. 1983. *Hen's Teeth and Horse's Toes*. New York: W. W. Norton.
- . 1989. *Wonderful Life: The Burgess Shale and the Nature of History*. New York: W. W. Norton.
- Hechter, Michael. 1987. *Principles of Group Solidarity*. Berkeley and Los Angeles: University of California Press.
- . 1992. "The Insufficiency of Game Theory for the Resolution of Real-World Collective Action Problems." *Rationality and Society* 4:33.
- Homans, George C. 1967. *The Nature of Social Science*. New York: Harcourt Brace & World.
- . 1968. *English Villagers of the Thirteenth Century*. New York: W. W. Norton.
- Kiser, Edgar. 1994. "Markets and Hierarchies in Early Modern Fiscal Systems: A Principal-Agent Analysis." *Politics and Society* 22:285-316.
- . 1996. "The Revival of Narrative in Historical Sociology. What Rational Choice Theory Contributes." *Politics and Society* 24:249-71.
- Kiser, Edgar, and Michael Hechter. 1991. "The Role of General Theory in Comparative-Historical Sociology." *American Journal of Sociology* 97:1-30.
- Kuhn, Thomas. 1970. *The Structure of Scientific Revolutions*, 2d ed. Chicago: University of Chicago Press.
- Kuran, Timur. 1995. *Private Truths, Public Lies: The Social Consequences of Preference Falsification*. Cambridge, Mass.: Harvard University Press.
- Lichbach, Mark Irving. 1995. *The Rebel's Dilemma*. Ann Arbor: University of Michigan Press.
- Lloyd, Christopher. 1993. *The Structures of History*. Oxford: Blackwell.
- Mason, Stephen F. 1962. *A History of the Sciences*. New York: Collier Books.
- Mokyr, Joel. 1993. "The New Economic History and the Industrial Revolution." Pp. 1-131 in *The British Industrial Revolution: An Economic Perspective*, edited by Joel Mokyr. Boulder, Colo.: Westview.
- Moore, Barrington, Jr. 1966. *Social Origins of Dictatorship and Democracy*. Boston: Beacon Press.
- Opp, Karl-Dieter, Peter Voss, and Christiane Gern. 1995. *Origins of a Spontaneous Revolution: East Germany, 1989*. Ann Arbor: University of Michigan Press.
- Reuschmeyer, Dietrich, Evelyn Huber Stephens, and John D. Stephens. 1992. *Capitalist Development and Democracy*. Chicago: University of Chicago Press.
- Schelling, Thomas. 1960. *The Strategy of Conflict*. London: Oxford University Press.
- Skocpol, Theda. 1979. *States and Social Revolutions*. Cambridge: Cambridge University Press.

- Skocpol, Theda, and Margaret R. Somers. 1980. "The Uses of Comparative History in Macrosocial Inquiry." *Comparative Studies in Society and History* 22:174–97.
- Somers, Margaret R. 1994. "Narrative and the Constitution of Identity: A Relational and Network Approach." *Theory and Society*.
- . 1996. "Where Is Sociology after the Historic Turn? Knowledge, Cultures, and Historical Epistemologies." In *The Historic Turn in the Social Sciences*, edited by Terence J. McDonald. Ann Arbor: University of Michigan Press
- Tilly, Charles. 1978 *From Mobilisation to Revolution*. Reading, Mass.: Addison-Wesley.
- Tullock, Gordon. 1971. "The Paradox of Revolution." *Public Choice* 1:89–99.
- Wrigley, E. A. 1988. *Continuity, Chance, and Change: The Character of the Industrial Revolution in England*. Cambridge: Cambridge University Press

Explanation in Historical Sociology: Narrative, General Theory, and Historically Specific Theory¹

Craig Calhoun
New York University

Sociologists seem doomed to fight the *methodenstreit* again and again. For every practitioner of intelligent division of labor or multidimensional research, there are a dozen advocates of one against the other.

Reopening the debate, Kiser and Hechter (1991) argue that historical sociologists have modeled themselves too much on historians, seduced by the desire to understand particular cases and the temptation to find knowledge through induction rather than general theory. In response, Skocpol (1994) denies that this argument sticks to her, and suggests that in any case it says little about specific explanations as distinct from abstract standards. Quadagno and Knapp (1992) argue that Kiser and Hechter have too narrow a notion of theory and that they fail to see how narrative history can help to build theory. Somers (in this issue) defends narrative as a dimension of theory itself. Along with the other critics, Somers also responds vigorously to the way in which Kiser and Hechter combine an ostensibly neutral, abstract argument for general theory with advocacy for rational choice theory. Some of Kiser and Hechter's position on general theory—such as the claim that “all good sociological explanations” must contain separate arguments about the motives of individual actors—thus derives more from the substance of rational choice theory than from methodological principles independent of specific theories. Somers accepts the argumentative gambit offered by Kiser and Hechter: to debate approaches to historical sociology as a replay of the *methodenstreit*, reworked in line with new trends in the philosophy of science.

This is unfortunate, though Kiser and Hechter invite it. They argue not simply about what will make for better sociological knowledge but also on the more philosophical terrain of “what adequate explanations must entail” (1991, p. 4). The result is to bring into often confusing combination three different kinds of argument: (1) about general epistemological and

¹ Direct correspondence to Craig Calhoun, Department of Sociology, New York University, 269 Mercer Street, Fourth Floor, New York, New York 10003.

scientific principles, (2) about the role of theory in historical sociology, and (3) about the merits or demerits of rational choice theory as an exemplar of general theory. I will have the least to say about the last of these, since I think the merits of rational choice theory are best debated in terms of the new knowledge produced by its use rather than in terms of abstract sermons about what constitutes a good sort of theory.

The heart of the argument is about what counts as a good explanation and how such explanations might be pursued. Kiser and Hechter emphasize the centrality of causal explanations based on deductions from general theories that include abstract models, covering laws, and mechanisms. They denigrate other approaches to scientific knowledge. In this they follow mainstream philosophers from a generation or two ago, but they also lay claim to the more recently fashionable label of "realist" philosophy of science. Somers challenges this with a broader and more complex reading of (and to some extent intervention in) contemporary debates in philosophy of science and with her own claim to be a realist.² Resort to philosophy, however, does not bring clarity to the exchange (even though clarity is one of the main goods philosophers of science ostensibly have on offer). This is largely because the colloquy between Kiser and Hechter and Som-

² Somers designates Kiser and Hechter's position "theoretical realism" and her own "relational realism." In response, Kiser and Hechter claim that there is a clear "standard usage" among philosophers as to the term "realist" and that it fits their usage. Readers should be wary of attempts to settle arguments by claiming to discern clear standard usages in other disciplines. Kiser and Hechter do not make much use of the philosophical realist who has probably had the most influential engagement with sociology, Bhaskar (see, esp. 1978, 1986). As their only recent authority on the position, they cite Bunge (1996). He is indeed a prolific advocate of realism, but far from suggesting a standard usage, he finds it necessary to describe his position as "critical realism" to distinguish it from the many other claimants to the rhetorically desirable realist label. Putnam (1987) refers to the "many faces of realism." Generic realism, for what it is worth, is a philosophical position rejecting both extreme forms of idealism ("the world is just my dream") and thoroughgoing empiricism (the inductivism that Popper [1979] called "the bucket theory of the mind", see also Miller 1985, pp. 101–17). Along with common sense, it holds that what is real is not limited to either sensory experience or our internally developed ideas. It acknowledges nonobservables as part of reality. The point in philosophy of science is to give theory (and beliefs and conjectures that may not quite add up to theories) standing alongside empirical inductions. At least in most versions, then, progress in knowledge is held to depend on both new observations and new thought about available observations. Thought (including theory) may play a role in producing useful observations; some thoughts (conjectures) are also corrigible by such observations, these are the ones most useful to science. Realism does not entail a Popperian belief in the progressive corrigibility of theories by empirical refutations, but the latter is consistent with and typical of a realist position. At the same time, realism suggests that the human world is always partly the product of our thought (and hence our culture); we cannot grasp it as a purely external reality (see Putnam 1987).

ers replicates divisions among philosophers, including notably that between "analytic" and both "Continental" and pragmatist approaches.³ Philosophers are very smart people, and their inability to find a way out of this long-standing division reflects, among other things, the enormous power that differing perspectives, criteria of judgment, and decisions as to what questions to ask have over the production of knowledge.

Here, I wish to clarify some of what is at stake for the relationship among history, theory, and sociological knowledge. I wish also to argue against the rather extreme "either/or" choice between history and theory, description and explanation, narrative and causal analysis that readers might draw from the debate as currently presented.

RAMPANT INDUCTIVISM IN HISTORICAL SOCIOLOGY?

According to Kiser and Hechter (in this issue, p. 786), "Whereas historical sociologists traditionally had insisted on distinguishing social science from history, now many sought to do away with the boundary between these disciplines." This boundary Kiser and Hechter see as that between idiographic and nomothetic approaches, induction and deduction, particular descriptions and general theories. This was always tendentious language, and presumed that a clear distinction could be drawn between the two kinds of inquiry. It grasps some differences between the disciplines, but it also distorts; certainly it is a caricature of what contemporary historians do to say that they seek only descriptions of historical particulars. It also implies objections to a wide range of sociological work, identifying as mainstream only a very specific kind of explanatory enterprise—and consigning not just narrative historical work but most qualitative sociology and much survey research to the descriptivist, idiographic, inductive periphery.

Kiser and Hechter draw their evidence for an antitheoretical turn from some abstract arguments against certain uses of theory rather than from

³ In the article published as part of this debate, Kiser and Hechter show themselves to be less naive readers of philosophy of science than Somers suggests, but they do appear to have limited their reading to publications that suited their tastes, and to have grasped less well the divisions among philosophers. In an early draft of their critique of Somers, Kiser and Hechter referenced Searle (1993) to support their contention that "the vast majority of American philosophy departments refused to endorse views of the sort that Somers espouses." The vast majority of American sociology departments refuse to endorse rational choice theory, but one doubts whether this persuades Kiser and Hechter that it is false. In any case, "American" is a key word in the quoted sentence. The "analytic" taste in philosophy is especially Anglo-Saxon. "Views of the sort that Somers espouses" are given more respect by philosophers elsewhere.

an examination of actual research. In fact, recent historical sociology draws on theory and frequently advances it. Consider Gould's (1995) use of network theory, Goldstone's (1991) of structural and demographic theory, and Goodwin's (1997) of theory of emotions as well as networks (to take only prominent authors whose names begin with G).

Since the inception of the discipline, historical sociologists have theorized historical patterns of continuity and change in both evolutionary and nonevolutionary models. Sociologists have also turned to history for data by which to test more limited theories and have developed theories partly through induction from historical scholarship. The critical discussions of theory that Kiser and Hechter cite actually come from a period of resurgence in historical sociology. Though important work earlier in the postwar period contributed to this resurgence (Smith 1991), the sub-field only gained critical mass and institutionalization in the 1970s and early 1980s. That it did so then reflected widespread theoretical concerns: to address social change and power relations; to restore agency to sociological analysis; to "denaturalize" social conditions by seeing paths that shaped them and ways in which they might have been otherwise. Reaction against functionalism, an attempt to overcome the marginalization of Marxism, and new readings of Weber all influenced this renewal of historical inquiry within sociology. So did a desire to pursue theory by pursuing history.

Historical sociologists recognized that extant theories of many important dimensions of modern society were based on implicit historical premises that needed to be reconsidered. Modernization theory, for example, had incorporated an account of early modern European history, which it then used to generate propositions for the study of social change in the rest of the world. It was not enough to test the propositions so generated (many of which did indeed prove dubious); it was important to reconsider the premises themselves on the basis of newer historical research. In a literal and important sense, then, the goal of much of the new historical sociology was to change the basic inductions that formed part of the core of sociological theories and provided the basis from which deductions were made. For this purpose, it was important that sociologists turn to history with an agenda that went beyond testing propositions derived from general theory. Of course, part of what enabled new historical research to generate new findings was change in analytic perspectives—such as greater concern for "history from below" and the role of power and conflict in shaping social change.

Explicitly or implicitly, this challenged some of the post-*methodenstreit* understandings of an older generation. Many had treated the social world simply as so many neutral instances of naturally occurring phenomena. For the newer historical sociologists, especially Marxists and Weberi-

ans, the contemporary social world was a specific configuration produced by historical struggles. In other words, the social world as it presented itself at one time was a particular, a singular conjuncture, as much as it was a cumulation of cases of more general processes. To say, for example, that it was predominantly capitalist was not only to indicate what theory would fit best with many replicable empirical data within it, but it was also to characterize the whole and to locate it in history. This was a theoretical claim.

Over time, historical sociology became domesticated as a style of research in American sociology (Calhoun 1997b). It lost its theoretical agenda (and the political agenda that shaped the interest of many in the 1970s and 1980s). Many sociologists turned to historical research, but it became increasingly hard to identify any common theory or approach to theory as dominant in or unifying of historical sociology. This is somewhat ironic, since much of the impetus for the rise of historical sociology came from theoretical issues and arguments, but it reflects the extent to which the legitimation of historical sociology as a subfield within the discipline was achieved by portraying it (misleadingly) as a method.

Rather than emphasizing sociology's substantive need for history—the need for social theory to be intrinsically historical—Skocpol, Tilly, and others argued that historical sociology should be accepted because it was or could be comparably rigorous to other forms of empirical sociology. Moreover, it was the only way to fill certain important gaps in empirical knowledge—for example, of rare events or phenomena that only happen over a long time span. Tilly (1984) emphasized the operationalization of quantitative sociological research and analytic methods for historical use; he engaged in important comparative studies, but he also warned against overreliance on a comparative strategy. Skocpol (1979; Skocpol and Somers 1980) placed a distinctive stress on comparison. While Skocpol's (1979) cases happened to be historical, her analytic method was less distinctively so. Indeed, her later analysis of the creation of state welfare programs in the United States was much more intrinsically historical (Skocpol 1992).

Emphasis on method obscured the importance of theory. Still, it is not entirely clear why Kiser and Hechter present their argument mainly as a challenge to historical sociology. In the first place, if they are right about the desirability of pursuing generalizable knowledge through identifying lawlike relationships and causal mechanisms, then this applies to all sociology (and, indeed, to all science). Second, historical sociology seems neither an extreme nor a typical case by which to make the argument for this version of general theory; symbolic interactionist fieldwork might qualify better as the former, survey research as the latter. Third, and most important, Kiser and Hechter do not address in a sustained way the specificity

of historical research or the place of historical reasoning in theory.⁴ Doing so would have made the connection to historical sociology clearer.

Kiser and Hechter (1991, pp. 1–2) are concerned mainly that “comparative-historical sociologists are tending to move from arguments against *specific* theories to arguments against *theory in general*” (emphasis in original). But they are not interested in “theory in general,” that is, in the range of intellectual products and orientations sociologists call theory, so much as in what they call “general theory.” By this they refer to formulations of explicit assumptions, abstract scope conditions, and models that include both lawlike statements of relations and causal mechanisms. Such theories are to be the basis for deducing potentially testable propositions, including causal explanations of events. Their generality inheres both in the lawlike statements of relationships (which are ideally universal and at least uniform—in the sense that similar causes always produce similar effects) and in the mechanisms that are abstract enough to be used in analyzing multiple concrete events.⁵ Preference for this sort of theory is

⁴ The illustration Kiser and Hechter (1991) offer, based on studies by Skocpol and Mann, considers variations in state autonomy as a question abstracted from historical context. Contrast this to studying the development of the modern state as a historically specific process (which indeed both Skocpol and Mann do).

⁵ Most of what Kiser and Hechter discuss has been a familiar part of texts and classes in theory construction for some time. The partial exception is their strong emphasis on causal mechanisms. Here they join in a vogue for reviving an idea historically linked to notions of strict causality, but using it in the context of more probabilistic models. Rational choice theory has played an important role in encouraging more emphasis on causal mechanisms alongside covering laws or causal relationships. Kiser and Hechter are actually somewhat vague about what is involved here, and their language shifts a bit between seeing the mechanisms as necessarily deduced from lawlike statements, and as simply being lawlike statements but of a different order from relationships of implication or correlation. One way of seeing what is at stake is to recognize that classical ideas of causation considered it as a relationship of force; all causes were agents exerting power over patients. This account ran into trouble in the 17th and 18th centuries, especially among empiricists. Failure to observe causes operating as external relations of force led to a variety of theoretical responses. Among these was Hume’s argument that what we observe is the spatiotemporal conjunction of events, not the necessary connections that previous theorists understood as causes. The idea of causality then became, for Hume, a matter of psychological imputation rather than an empirical observation. The notion of mechanisms came to prominence as an attempt to identify the ways in which connections were actually formed between objects and is especially important to “realist” theories willing to discuss causation even when it cannot be directly observed. Such mechanisms come in two forms: actual energy transfers and triggers or signals that do not involve substantial material force. Many important causes in social life are of the latter sort, which means that there can be no presumption of a “proportionate” material relationship between cause and effect (Bunge 1996, p. 31). This is one reason why social life is full of nonlinear relationships (see Abbott 1988). These do not invalidate causal analysis, but they make it hard to use some traditional ways of reasoning about causation.

the basis for Kiser and Hechter's (1991, p. 2) complaint that the "new wave" of historical sociologists is too inductive: "The structure of their arguments . . . tends not to be implicative (involving deductive logic), but conjunctive (involving the use of coherent narrative)."

Behind this preference lie two distinct though related philosophical controversies. One is the classic question of the *methodenstreit*, whether the human sciences need be distinguished methodologically from the physical and biological sciences. Arguments for such a distinction are generally grounded in three claims. First, human nature is never fully independent from culture and is modified by historical change. Second, we do not understand human action adequately by explaining it solely in terms of external antecedent causal conditions; such factors need to be supplemented at least by inquiry into motives and purposes and the pursuit of empathic understanding (*verstehen*). Third, many of the important events concerning humanity happen only once or very rarely and thus cannot be explained by theories focused on patterns in repetitious events such as those of nature.

The second controversy concerns the problem of induction. Kiser and Hechter use the same term to designate both the belief that basic categories of thought (like causality) derive from psychological experience and the attempt to make empirical generalizations basic premises of social theory. The former is inductivism in a strict epistemological sense, the latter a more general sort of empiricism. Arguments against the former are not necessarily decisive against the latter.

The status of causal reasoning is a classic issue that has bedeviled philosophers since Hume. Hume ([1748] 1902) famously argued that all pure empirical induction could yield was the knowledge that events were conjoined in space and time. The idea of cause could not be given by experience alone. Hume was content simply to focus on empirical generalizations without reliance on the notion of an unobservable causation. Defending causality and the greater certainty of lawlike knowledge, idealists were prepared to argue then that an idea of causation was given to the human mind a priori. They, more than empiricists, are ancestors of today's "realists." Durkheim ([1912] 1965) famously accepted the argument that individual experience could not produce basic categories of thought, but held that causality and the others were not simply a priori, but given by society. Eventually, Popper ([1934] 1959, 1972) and others saw that part of the intractability of the problem lay in the very empiricist presuppositions with which Hume and others began. In Popper's phrase, they worked with a "bucket theory of mind," believing that there is nothing in our intellect that has not entered it through the senses.

Popper argued that we should distinguish psychological questions (e.g.,

about how we learn) from logical questions (such as whether an inference is valid). As a logical question, the problem of induction concerns centrally how we can determine the truth or falsity "of universal laws relative to some 'given' test statements" (Popper 1972, p. 8).⁶ Test statements are hypothetical singular descriptions of observable events; they may be judged true or false on the basis of experience (in science, usually experience guided by observational technique). They establish good reasons for accepting universal laws as true, though of course the latter remain conjectures subject to possible future refutation.⁷ Whether empirical conjunctures suggest causal connections to us remains a psychological matter; logically they are important as tests for propositions derived from conjectural laws, not the basis for such hypothetical statements. We may gain a great deal of knowledge from empirical generalization thus, but in itself this will not enable us to arrive at universal laws.

Kiser and Hechter maintain Popper's sharp distinction between the context of discovery and the context of justification, but are somewhat confusing in this regard. They purport to describe, for example, "the failure of induction to suggest causal mechanisms" (1991, p. 15). This is an extreme overstatement, however, since many causal mechanisms have been suggested to scientists of all sorts (including historical sociologists) by direct inductive observation (combined with some reasoning). Kiser and Hechter are right, however, that causal mechanisms cannot be specified on the basis of induction alone. Because they are not directly observable, their specification depends on inference from indirect tests. It is the latter in which Popperian scientists are interested. It does not matter to them so much how social scientists get their ideas as how they test them. Kiser and Hechter might reply that they would prefer that explanatory ideas be deductions from general theories, but it is unlikely that all new ideas as to causal mechanisms will be produced this way (even if one thinks this is the best way to formulate them). It would be more accurate to say that for the view they express it does not matter whether scientists get their ideas by deduction or by inspiration while taking a shower. What matters is that they are able to account for those ideas in terms of the deductive structures of general theories so that their possible falsification can contribute to the advance of theory.

⁶ "Given" is in inverted commas because Popper wishes to call attention to a further class of problems pertaining to the empirical validity of test statements, including whether we can speak of experienced instances prior to the theories that define them as instances of test statements

⁷ This is one of the key distinctions between Popper's critical realism and positivism. Positivists commonly hold that certain truths are established outside this process of infinite corrigibility.

By contrast, Somers (p. 732), in line with Kuhn and recent science studies, would argue that the idea of a sharp separation between contexts of discovery and justification is misleading. It allows for scientists to represent their work as fitting more fully into canonical models of inquiry than it usually does, and it systematically underplays the concrete histories in which new data, understandings, and theories are produced. If in fact changes in theories are always historically produced in a field of available alternatives, then formal justifications in terms of covering law theories and the like are at best partial accounts of why the changes have occurred. One must account also for what has been studied and what has not, for what sorts of observations could be gathered under given research conditions, and so on. Which topics are considered interesting, which archives are open, what technologies aid access? There is no reason to assume that these are simply uniform filters on progress; they shape sharp turns in the orientation of knowledge as well as variations in its rate of accumulation. Accounts of such factors may take narrative form. It is worth noting, though, that similar strictures apply to narrative. It is not sufficient simply to tell the story of what has occurred without inquiring into the conditions of these empirical events and the factors that selected this actual history from other possible ones, and into the rhetorical tastes and practices that guided the construction of the narrative.

GENERAL THEORY AND HISTORICAL EXPLANATION

When historical sociologists speak of induction, they seldom mean to enter into debates on basic epistemological problems. Proponents usually mean, rather, to suggest that the amount of insight or "explanation" provided by highly general theories is modest, partly because such theories do too much violence to the specifics of the historical patterns. These proponents propose to build knowledge of those patterns "upward" from empirical specifics.⁸ Opponents, like Kiser and Hechter (1991, p. 9), charge that this

⁸ As it happens, many of the "inductivist" historical sociologists rely heavily on the published works of historians to produce the empirical descriptions from which they work. The issue is not precisely nearness to archives, nor even nearness to facts, but something more like breadth and depth of knowledge of conditions and events. Thus Stinchcombe's (1978) call for historical sociologists to pay less attention to general theories and more to analogies between empirical instances is a call for attention to complexity of patterns in richly understood cases, and comparisons among such cases, rather than fitting of such cases into abstract theories. But note here that Stinchcombe works with a sense of general theory different from that held by Kiser and Hechter. The theories to which he (and many of the so-called inductivists in historical sociology) object are theories of "whole events" or even whole historical processes. We may learn less from a general theory of revolutions, thus, than from detailed comparison of a

is a sign of "the predominance of historians' norms" in historical sociology, with the result that "scope (generality) and analytic power have been minimized and descriptive accuracy has become the predominant criterion for constructing and judging explanations." Their complaint, in other words, is less that historical sociologists are attempting wrongly to produce general theories by means of induction than that they are dispensing with general theory in favor of empirical detail.

This empiricism, it would seem, is what Kiser and Hechter mean by acting too much like historians. It is indeed a stylistic quirk of historians to point to the particulars of cases they know well in order to challenge attempts at generalizations; this makes them threatening book reviewers for historical sociologists (see Kiser and Hechter 1991, pp. 23–24). Inspired by 19th-century German "historical science," American and some other historians took on the project of studying the past "as it really happened" (Novick 1988). This became a hallmark of objectivism and a challenge to all who would read the concerns, values, and indeed theories of a later age into historical accounts. This was central, for example, to what E. P. Thompson meant when he proposed "to rescue the poor stockinger, the Luddite cropper, the 'obsolete' hand-loom weaver, the 'utopian' artisan, and even the deluded follower of Joanna Southcott, from the enormous condescension of posterity" (1968, p. 13) and also when he later railed against "the poverty of theory" (1978). But despite an ideology of "nothing but the facts," historians seldom engage in radical inductivism. They do not simply walk arbitrarily into archives and demand documents, storing up facts as the experience dictates. They seek to test existing claims and extend understanding to new (or hitherto unstudied) events. They treat studies of specific towns and social movements, fertility rates and royal succession crises as cases to be compared as well as appreciated in their specificity. They have theories, even when they choose to express them implicitly rather than explicitly. Above all, in light of the present debate, they write narratives that are hardly exercises in induction or indeed in pure empiricism.

Whatever theory historians use to organize their narrative explanations, however, is commonly left implicit. Historical sociologists have related to history partly by making such theory more explicit, subjecting it to logical

handful of revolutions (see Skocpol 1978). But this is a claim about the "context of discovery" rather than the "context of justification." In urging us to seek analogies, Stinchcombe might be understood to be suggesting a procedure well suited for identifying causal mechanisms, as Kiser and Hechter would like, rather than arguing against these. But like Robert Merton (1968) in advocating middle range theory, he is indicating that we are at an early stage of our quest and ought not to leap to premature attempts to theorize complex wholes

scrutiny, and challenging it substantively. Conversely, they have also related to theory by challenging its historical presuppositions. As I suggested at the outset, and as Somers (p. 731) argues, "all social and political theory is founded on presuppositional historical claims." Though these may figure in the "core" of the theory (in the sense of Lakatos [1978]) and thus not be rendered as propositions subject to empirical test, they may be challenged directly as to their plausibility. Argument on such points is appropriately based on historical research.

When Kiser and Hechter refer to explanation, they take it for granted that this means causal explanation in a covering law model.⁹ At least for the purposes of social science, they do not consider it an explanation to say what something means, or what its social function is, or by what sequence of events it came to exist or take on its present form. The kind of causal explanations that interest Kiser and Hechter are explanations for the occurrence of "facts," understood as highly replicable observations generated under specific theoretically defined conditions (see Bunge 1996, p. 137). In adopting the covering law model for explanation, they necessarily side with the position that there is no deep difference between the human and the natural or physical sciences. One side in the *methodenstreit* argued that historical change, cultural difference, and individual action made such a difference in human affairs that explanation could not proceed in the same form as it did in physical sciences or even in biology (though biology was understood to be more like history and social science than, say, physics or chemistry were). The basic idea of explanation through covering laws is that an event is explained by subsuming it under universal laws. More specifically, an explanation requires both statements of conditions (such as scope of application) and laws (which, as we saw in discussing Popper, take the form of well grounded but still always contingent hypotheses). The structure of a covering law explanation is virtually identical with that of a prediction.¹⁰ Not everything in the theory is directly testable, of course, only the empirical statements that are deduced from it. Its core of assumptions, including basic mechanisms and models, can only be tested indirectly through tests of the propositions they yield.¹¹

⁹ The covering law model has many adherents who differ on other specifics. It is linked especially with Popper, Hempel, Mandelbaum, and Nagel, each of whom has written specifically about its application in historical explanation. Hempel (1942) wrote perhaps the decisive article in the ascendancy of the covering law model in the philosophy of history.

¹⁰ See Bunge (1996, chap. 5) for a discussion of the relationship between explanation and prediction, including reasons why they are not precisely identical.

¹¹ There is a substantial literature concerning what counts as testing and how one may (or may not) evaluate "progress" in scientific theories. Lakatos (1978) introduced the idea of a "progressive" scientific research program to describe the capacity of a given

Kiser and Hechter (especially in the discussion published here) associate deviation from covering law explanations with the recent rise of "relativist" science studies, but in fact dissent is long standing. There have been sympathetic critics who held that the standard of universality set up by the protagonists of covering law explanations was much too stringent. A sharper challenge has come from adherents of varying forms of narrative explanation (Ricoeur 1984). Thinking of natural more than human history, Goudge (1958) argued that historical explanations commonly, and appropriately, rely not on laws but on the specification of a number of conditions that are jointly sufficient to account for an event when combined in a coherent narrative sequence. Gallie (1968) maintained that even joint sufficiency was too strong a criterion; it is enough that the temporally prior necessary conditions be specified. Rather than deducing a class of similar events from covering laws and instantial conditions, thus, one explains historical events by placing them in a sequence of statements of conditions. This is the distinction between "implication" and "conjunction" to which Kiser and Hechter allude. Lest it be thought that "conjunctive" explanation is mere recourse to Humean association between temporally contiguous events, note that the criterion of narrative coherence sets quite different standards, even though they are not of the covering law sort. Stories come in different genres and a range of disciplines and constraints shape the formulation of narratives—explanatory or otherwise. Danto (1965), for example, holds that the involvement of a subject in a story is crucial to narrative explanation. In any case, a story is neither a regression of variables against time nor a mere record of events. The identification of a causal sequence "intrudes" into narrative to give it direction (Gallie 1968, p. 110).

Somers enters the fray with a more substantive defense of narrative and a further challenge to covering law explanation. Before proceeding to this, however, we would do well to be a bit more probing about what covering law explanation and the use of general theory might mean in actual historical sociology.

Kiser and Hechter are less than clear on this point partly because they do not distinguish adequately between two claims about the importance of general theory in historical sociology. One is that general theory is valuable as the source of explanations in historical sociology because it allows us to understand better whatever substantive topic of inquiry motivated us in the first place. The examples Kiser and Hechter give generally pertain to this idea of value. Their major claim about rational choice theory

theory to continue to guide research that produces progressively more successful empirical explanations, while suffering only modest revision at some distance from its core.

is that it will be useful to historical sociologists, not that it needs historical sociology to advance as a theory. The other claim, however, follows even more basically from the covering law approach to science. This is that general theory should be employed in historical sociology not so much to advance immediate historical understanding (though we hope this will happen) as to advance general theory itself by bringing it under risk of refutation and stimulation of further creativity by distinctive empirical data.

In this connection, it is interesting to note the critique of historical sociology offered by John Goldthorpe (1991), almost simultaneously to that of Kiser and Hechter's first foray (and surprisingly not mentioned in the present debate by either them or by Somers). Goldthorpe is also an adherent to a covering law approach to explanation and to the idea that the development of good general theory is the main object of sociological research. He agrees with Kiser and Hechter that many historical sociologists are inappropriately blurring the boundary between history and sociology (somewhat confusingly taking Skocpol as a primary example, though she has been a strong defender of the disciplinary distinction). Despite the similar premises, however, Goldthorpe reaches a more or less opposite conclusion. The burden of his argument is that sociologists have little business doing historical research. If they do it well they will simply be historians, though the odds of this are small given sociologists' poor training for doing historical research. More important, historical research is poorly suited to generating the observations needed to advance general theory by providing empirical refutations or confirmations for its propositions. What is needed for the development of more or less universal, lawlike theoretical statements is a large number of replicable observations, preferably gathered under controlled conditions. According to Goldthorpe, historical facts are inferences from relics, while the preferred facts of social science are the results of new, more perspicuous, more complete and repeatable observations. "History may serve as, so to speak, a 'residual category' for sociology, marking the point at which sociologists, in invoking 'history,' thereby curb their impulse to generalize or, in other words, to explain sociologically, and accept the role of the specific and of the contingent as framing—that is, as providing both the setting and the limits—of their own analyses" (Goldthorpe 1991, p. 14). History is more about scope conditions than causal generalization, thus. The more we seek the latter, the better off we are gathering contemporary data.

Why then turn to historical inquiry? Kiser and Hechter give examples of recent historical sociology in which general theory (rational choice) is employed to good effect, but these do not answer the question. That is, Kiser and Hechter give us no reason to think that the studies were decisive for advancing the general theory as such—for example, for developing

new micromechanisms that would change analysts' views of choice processes.¹² Rather, rational choice theory is shown to be useful for answering certain particular questions about events and phenomena that are of interest for other reasons. One might be interested in certain historical events or processes for their own sake, or one might care about a specific historical narrative—say the history of one's own nation—which would be advanced by a better understanding of certain events or processes. Presumably, however, these sound too much like historians' reasons for Kiser and Hechter's taste. This is basically where Kiser and Hechter leave us: rational choice theory is argued to be useful in analyzing problems the interest and importance of which it cannot in itself explain.

Alternatively, one might be interested in history because one wished to advance theory that was itself intrinsically historical. That is, one might be interested in a better theory of capitalism and thus understand capitalism not simply as recurrent events and processes but as a historically specific and situated phenomenon that not only shapes an epoch but is shaped by other historical changes and by phases in its development. In any everyday sense, this would seem an extremely general theoretical project, but it is not a project of developing general theory in Kiser and Hechter's sense. Their sense identifies generality with replicability; that is, specifying events or processes that will occur over and over again in many cases when causal conditions are met. This latter approach allows for the study of many processes within capitalism—investment practices, inflation, labor regimes, and technological obsolescence. It does not allow for the study of historical capitalism as such.

Theory that is intrinsically historical is not quite theory in Kiser and Hechter's sense. Marx's *Capital*, Durkheim's *Division of Labor*, and Weber's *Protestant Ethic* are among the many works that do not qualify. Of course, instead of just dropping them from the canon of sociological theory, we could reanalyze each, trying to find within it a residuum of useful propositions that could be recast in an ahistorical, covering law

¹² Kiser and Hechter (p. 797 above) do assert that translating historical scope conditions into abstract ones allows for insights developed through inquiry into the past to be applicable to the present and future, and indeed that this is "the most important contribution that historical sociology can make to the discipline as a whole." This does not answer the question of why one would need to turn to historical cases to gain these insights. It also seems to underestimate the complexity of the task of translating historical into abstract scope conditions and the extent to which choice of theoretical languages is likely to influence this. Most important, this presumes the capacity to abstract cases altogether out of history. This is most plausible when one is concerned with frequently replicated phenomena, it completely misses problems of historical singularities (like the rise of capitalism) significant for an entire epoch. Of course, the frequently replicated phenomena may be parts of more complex and rarer large-scale phenomena and may shed light on them

manner.¹³ Nonetheless, this would both do violence to these major works and eliminate much of what has given them enduring theoretical importance.

My point is not to dismiss the approach to theory that Kiser and Hechter advocate—either in general metatheoretical terms or in the specific example of rational choice theory. Rather, I want to suggest that it is simply not broad enough to encompass the full range of meanings of theory, the different ways in which theory is important to sociological understanding, or the reasons why sociologists turn to theory. It is a useful part of the disciplinary tool kit, better for some kinds of intellectual projects and problems than others, but poorly understood as definitive of theory (or science or the advancement of knowledge) in general.

NARRATIVE AND HISTORICAL EPISTEMOLOGY

Somers challenges Kiser and Hechter on many fronts and raises a variety of concerns. Centrally, she argues that theory cannot escape history and both misunderstands itself and introduces potential distortions into its work if it tries to do so. The kind of general theory that Kiser and Hechter advocate is, as we have seen, basically ahistorical (or, to put it more positively, omnitemporal). This has two implications. One is that, while this sort of theory can contribute to historical explanations, it does not address the specifically historical dimensions of theory or explanatory problems. The other is that theory of this sort tends to leave unexamined the “presuppositional historical claims” that inform it (Somers, p. 731). One of the points of Somers’s historical epistemology is to keep attention focused on the embeddedness of all knowledge (as of all social life) in history. This implies continually reexamining the ways in which both the standpoint of the present and the specific histories of theory development shape what we ask and do not ask and therefore what we know and do not know.¹⁴

¹³ Elster (1985) has done something of this in the case of Marx. For a contrasting view, in which Marx appears as a historical and critical theorist rather than an immature rational choice theorist, see Postone (1993).

¹⁴ There is a close link here to arguments about the inescapability of hermeneutic interpretation and the necessary limits to covering law theories and scientific prediction. As Taylor’s ([1971] 1985, p. 48) classic statement goes, man is a “self-defining animal.” “With changes in his self definition go changes in what man is, such that he has to be understood in different terms. But the conceptual mutations in human history can and frequently do produce conceptual webs which are incommensurable, that is, where the terms can’t be defined in relation to a common stratum of expressions.” Elsewhere (Calhoun 1995) I have argued that such cultural incommensurability is better understood as involving practices than simply terms. The very process of creating our contemporary webs of meaning—including our theories—implicates us in some rather than other relationships to projects of historical understanding.

Somers's argument is rooted in a crucial point derived from Kuhn: the history of the development of science explains as much about its current patterns of knowledge as do its attempts at theoretical justification of its knowledge claims. Theories—or paradigms—are seldom falsified (let alone proved true). Rather, scientists are persuaded to shift from one theory to another available one.¹⁵ This does not mean that there is no such thing as falsification, rather that only fairly narrow empirical claims get falsified. Falsification of these claims casts doubts on underlying theoretical structures from which they may be deduced or to which they may be important. However, this is seldom if ever decisive in the actual abandonment of theory. Most reasonably robust theories are indefinitely revisable, and in any case an alternative theory needs to be presented that can succeed in competition with the one previously held. This may be a new theory or, as often happens in social science, the recycling of an old, previously widely abandoned theory, now presented in modified form. This last is precisely what happened with the redeployment of rational choice theory in sociology (no longer called utilitarianism and shorn of some of the assumptions that would have tied it to that early incarnation). Functionalism has likewise come back into vogue after deep crisis. Marxism may rise again, and so forth.

The significance of these points is more far-reaching than at first appears. For example, they raise questions about the idea of scientific progress. There are clearly new achievements in scientific understanding and in technological capacities to manipulate the external world that are derived from science. No serious scholar doubts this, and the charge of relativism hurled at post-Kuhnian science studies is at least overstated if not a red herring.¹⁶ The problem is not that science does not help people to do new things. It is rather that it is hard if not impossible to arrive at a defensible metatheoretical position from which to assert that the succession of theoretical understandings represents a linear "progress." A commonly stated ideal is that a new theory should be able not only to succeed in all the explanatory tasks where a previous theory succeeded, but also to succeed where the old theory failed and perhaps even explain why the old theory failed. Even in the paradigm case of relativity theory versus

¹⁵ In the following discussion, in accord with widespread sociological usage, I will generally use the word "theories" for the complex structures of theories and research approaches that Kuhn termed paradigms and not attempt to restrict it to a narrower sense of specific theories within a broader paradigm.

¹⁶ In a typical overstatement, Kiser and Hechter (p. 788) accuse Somers of holding that "science is nothing but a social construction," as though she implied that all social constructions deserved equal epistemological credence. In claiming that science is a social construction, she never claims that there is nothing more to say about it or its efficacy, nor does she take a radically relativist (as distinct from relational) position.

Newtonian mechanics, it is not clear that this is a good account of what happened. In any case—and this is important to Somers's argument—it is a narrative account of what happened, not a set of changes in understanding deducible from general laws of either nature or human reasoning.¹⁷

When theories contend and one gains more widespread adherence, this is seldom based on a confrontation between each of the theories in its entirety. Rather, certain parts or features of each theory are brought to the fore and become crucial to the assessment of their relative explanatory power. Such assessment takes place, moreover, in relation to specific explanatory problems. These problems are produced by research projects that direct attention in some possible directions rather than others and that focus attention on certain aspects of phenomena and not others. It is often in the choice and framing of research problems that extrascientific forces (e.g., funding agencies, governments, employers) have their greatest effects. Extrascientific efforts to promote particular solutions to problems are likely to be much less influential in the long run—at least where relatively open scientific practice is involved. There are also reasons within science but outside theory that help determine which problems become the foci of research and the basis for contention among theories. For example, available research technologies and the distribution of experience with different material practices among scientists and laboratories may shape what scientists choose to study. Scientific knowledge comes not simply out of theories and tests but out of the "mangle of practice" (Pickering 1997). There is a potentially infinite range of research problems, so the confrontation between theories is played out on an arbitrary subset. The subset may be historically explicable but is unlikely to be amenable to satisfactory justification within any theory let alone across theories. Thus which problems scientists study (and which they do not) has an impact on the direction of theoretical development. Even where there is clear improvement in the ability to solve the particular problems addressed, we have no basis for assessing this in relation to the problems not addressed.¹⁸

For these (and related) reasons, scholars in science studies since Kuhn have argued that it is important to study the context of discovery as well as efforts to justify knowledge claims. The context of discovery determines

¹⁷ Kuhn ([1962] 1970) did suggest that "rational reconstruction" could save the narrative history of change in scientific theories from being subject only to hermeneutic interpretation. Denying this is basic to both a more thoroughgoing pragmatism (Rorty) and many so-called postmodernist challenges (e.g., Lyotard).

¹⁸ This holds without accepting Somers's much stronger (and more contentious) statement that "there are no universally valid principles of logical reasoning; there are only problem-driven ones" (p. 766).

both the range of theories available for comparison at any one time and the specific research practices through which empirical observations may be generated. Part of what Somers challenges in Kiser and Hechter is their adherence to a discussion of theory and theory testing based mainly on justification and their neglect of the significance of this dimension of Kuhn's argument.¹⁹

In a closely related point, Somers emphasizes Kuhn's argument that it is much more difficult than most philosophy (and ideology) of science suggests to "articulate" or develop "points of contact between a theory and nature" (Kuhn 1970, p. 30; quoted by Somers, p. 735 above). It is conceptually as well as sometimes materially hard to generate observations that are genuinely effective in deciding for one theory and against another.²⁰ As Somers points out, neither induction nor deduction solves the problem. What exists between theory and research practice is not a relationship of simple testing but one of sustained mutual engagement.

CAUSALITY

Kiser and Hechter (1991, p. 4) begin, Somers stresses, with a dubious empirical claim. They assert that there is "wide agreement . . . across social science" that causality is "the first requirement of an adequate explanation." Somers rightly responds that no such agreement exists, though I think we might qualify this by noting that many social scientists, perhaps even a majority, think casually that explanation is basically a matter of identifying causes. The problem is that they do not go beyond such casual thinking or they would run into the many difficulties with the assertion. Somers perhaps does not distinguish clearly enough three kinds of difficulties: (1) Different things are meant by explanation. (2) Different things are meant by causation. (3) Even where social scientists definitely define explanation as adducing causes, and mean the same thing by causes, such causes are notoriously hard to prove.

¹⁹ I think this is also what Quadagno and Knapp (1992, p. 482) are getting at when they write that it makes sense to "derive theoretical explanations from detailed narrative histories." Kiser and Hechter (pp. 795–96) mock them for asserting this, and for asserting that theory should provide researchers with questions, not just with answers to be tested. Obviously, Quadagno and Knapp cannot mean the same logical operation by "derive" as do Kiser and Hechter or anyone else speaking about deductive theorizing. But they are within the bounds of reasonable English usage, and they call attention to the difficulty of making sense of the role of theory without addressing the practical production of knowledge, the context of discovery.

²⁰ The issue is partly sheer difficulty, partly the tendency to be biased by theoretical starting point. Somers (p. 760) approvingly cites Hollis (1994, p. 79): "In deciding on which facts to select and transform into evidence, we are already to a large degree deciding between rival theories."

Explanation.—The equation of explanation with causality has indeed become widespread and is especially widespread in Anglo-Saxon philosophy of science and scientific discourse. It is less widely shared in other philosophical traditions and in ordinary language. Many social scientists are influenced by other philosophical traditions, and quite a few even speak ordinary language. Some, therefore, believe they can “explain” something by saying what it means or how it works, not just what caused it (or what it causes). This view is particularly salient in relation to historical inquiries (as to cross-cultural ones) because considerable effort may be required to grasp the meanings that objects, phenomena, ideas, and other people have for actors in very different social and cultural settings (see Calhoun 1995, chaps. 2–3). As terms of art, scholars sometimes distinguish achieving understanding, in the sense just evoked, from explaining. This follows the distinction between *verstehen* and *erklärung* formulated during the *methodenstreit*. The point of the distinction was to contrast the attempt to produce knowledge through interpreting how historical actors thought and felt, their practical orientations as well as their explicitly avowed intentions, and the attempt to produce knowledge through study of external causal linkages. Understanding was held to be important (by some) because the human sciences dealt with sentient beings who took actions based on the meanings they found in or imputed to their worlds. Explanation was held to be different not because it was more empirical, but because it was more objective—including in the literal sense of treating people as objects rather than actors (in other words, making no distinction between action and behavior). This notion of causal explanation achieved its apogee in social science (especially psychology and sociology) in the era of behaviorism.

Covering laws and mechanisms are both examples of explaining by adducing external factors to account for behavior (or other events). The interesting, perhaps even ironic, twist introduced by Kiser and Hechter (along with Elster [1985], Coleman [1984], and many others) is to emphasize causal explanation in an ostensibly voluntaristic theory. At least in many versions, rational choice theory treats human beings as actors with some degree of agency and makes reference to their motives and not merely the external causes evoked by those who linked the notion of cause with that of explanation during and after the *methodenstreit*.²¹ On Som-

²¹ Bentham was a strict behaviorist. This is a difference between some versions of modern rational choice theory and classical utilitarianism. Rational choice theory in this sense tries to combine what philosophers have sometimes distinguished as a logic of natural relations and a logic of action (Dray 1964). Kiser and Hechter (pp. 789, 800) specifically identify their theory with intentional action rather than a more deterministic logic of interests and argue that because of the importance of individual action, sociologists “must employ a different methodology than that used in the natural

ers's reading (pp. 749, 752 above), Kiser and Hechter offer a theory of "agents as mechanisms."²² In doing so, they not only introduce rather simplistic ontological claims about what people are like; more surprisingly, they withdraw the apparent voluntarism of the theory. This is so, she suggests, because they hold that mechanisms must be deducible from general theories, so if an agent is a mechanism she cannot have much real agency. The demand for microfoundations as part of all fully successful theoretical explanations starts out looking like an opening to human reason (as in the *verstehen* side of the *methodenstreit*) but (if the view of the agent is as narrow as Somers suggests) ends up introducing a substantial determinism and possibly an element of tautology into the theory.

Thinking of explanation only in terms of causality, and of theory only in terms of the general theories that generate causal explanations, leaves out many important efforts to achieve relatively general understandings of social life. I have alluded already to the reasons why one might wish to formulate intrinsically historical theories. A further example may be clarifying specifically in relationship to causality. Some abstract categories are important in theory, research, and everyday social life because they constitute dimensions of the reality we study and in which we live. Money is such a constitutive abstraction; its creation is a collectively achieved performative investiture of value. The business corporation exists and is treated in law as an artificial person (and so, in many respects, is the state; see Coleman 1984). Nations are not simply found in external terms but created in human history (though not necessarily through such discrete events as the idea of "invention" implies). Their creation is partly based

sciences." They do not say what this different methodology is; however, though they do imply that it remains objectivist, rather than being strongly interpretative. They maintain a strong distinction between individuals and contexts (p. 800) and insist on "a social world existing outside the minds and sense perceptions of its participants" (p. 789), and thus apparently not intersubjectively constructed.

²² Somers changes her account somewhat on p. 764 when she endorses Tilly's (1995, p. 1602) account of rational choice theory's treatment of actors as "coherent, durable, self-propelling social units-monads." At this point she is interested (reasonably) in emphasizing the lack of a strong sense of the agents as social. She shifts, however, from describing the agents as being mechanisms to describing them as the bearers of mechanisms. In either case, they are causal forces in themselves, but in the second case one could not conclude that Kiser and Hechter had taken away apparent voluntarism as clearly as one could in the former. A somewhat related issue arises a page later. Somers (p. 764) accuses Kiser and Hechter of "essentialism." It seems to me that they (and most rational choice theorists) have not actually offered a substantial enough account of how they understand human agents for us to decide on this charge. It may be, however, that the only way for rational choice theory to escape the charge of essentialism is to relax its standards for microfoundations. If the agents are indeed mechanisms, or have inherent causal power by virtue of fixed dispositions, then the charge of essentialism is more likely to stick.

on a rhetoric or discursive formation that makes it meaningful for people to speak and think about and become emotionally committed to "their" nations (Calhoun 1997a). To paraphrase W. I. Thomas, each of these phenomena becomes real because we treat it as real in our actions. The existence of money, business corporations, and nations is basic to the modern world. Explaining how this is so would seem to be important for sociological theory.²³ But is this project of explanation graspable entirely through covering laws and causal mechanisms? Part of what we want to understand is the transformation of the world by these new phenomena that are created in and through culture. Explanation here may be achieved partly through narrative, as Somers suggests, but this does not exhaust it either. Critical inquiry into cultural categories may produce theoretical explanations that are neither simply causal nor simply narrative.²⁴

Causation.—Even in the heyday of covering law theory, one of its foremost defenders could write of the principle of causality that "there is no generally accepted standard formulation of it, nor is there general agreement as to what it affirms" (Nagel 1961, p. 316). Most of what passes for causal analysis in the social sciences is in fact identification of more or less "weak implication" between statistical variables (Boudon 1974). It is very different from arguments as to strict causation. The issue here is not specific to history; in fact, highly particularistic historical arguments may sometimes come closer to the logic of strict causality than most quantitative social science. Somers (pp. 746–47; original emphasis) implies that it is an error for Kiser and Hechter to "attempt to combine a covering law model logic based on *empirical* regularities with a method of *imputed*

²³ As Somers notes, Kiser and Hechter make a good deal of the inability of comparative historical sociology to achieve independence of cases. Their solution is to turn to lower level, frequently repeated phenomena where they believe this is less of a problem. Somers (pp. 757–58) accuses them of presupposing (if not quite of holding) "that cases are discovered as ready made, discreet entities—given in the nature of things, that is—rather than constructed along analytic parameters." While this may not be altogether fair, it is certainly true that Kiser and Hechter offer little account of how one would make the construction of cases an object of study, facilitating inquiry not just into theory and its operationalization, but into the way the categories used by many theories reflect the historical constitution of aspects of the world they observe (as, e.g., by assuming nations, or events situated in national histories, as the units of comparison).

²⁴ One may argue that the first chapter of *Capital*, vol. 1 (Marx 1867), is a theoretical account of such constitutive categories, of how the commodity form and abstract labor constitute capitalism. This is not per se an account of the historical origins of capitalism—the sequence of steps by which it came into being. Nor is it a covering law theory of either capitalism or the various features Marx identifies as central. Nor further is it an account of the commodity form or abstract labor as causal mechanisms producing capitalism. Rather, it is a theoretical statement of what capital (the basic term in capitalism) is. In this sense, but not in Kiser and Hechter's, it is an explanation.

causal mechanisms.²³ This is not necessarily so. It is true that positivists attempted to dispense with reference to unobservable causal mechanisms, but not that all covering law theorists follow them in this. Moreover, until the 1960s, most covering law theories were theories of strict causation. Nagel (1961) and Hempel (1962) led the way in accepting probabilistic and statistical laws in place of universal ones. An important result of this, however, doesn't get the attention it deserves from Kiser and Hechter. This is the fact that when "laws" are probabilistic, it is no longer possible in the ordinary logical sense to deduce an explanandum from an explanans, though one may infer probability.

Problems of proof.—Proving causation is a formidable challenge; it is met to varying degrees of satisfaction often enough that it remains meaningful to speak of causal analysis in social science, but it is met in such different ways that care is required before assuming common meaning. Some will accept a causal argument as complete if it consists of necessary but insufficient causes, others require necessity plus sufficiency. Some hold that causality can link only events, others use the term to describe links between properties or conditions and events. There is also the difficulty that most phenomena in social life admit of multiple causes (under any meaning of the term). Identical events, moreover, may result from completely different causal determinations (though some still dispute this, sticking perhaps to Comte's positivist dictum that there should be a one-to-one match between causes and effects). Not least of all, there is the question of determinism. Indeterminacy in the form of chance operates alongside causality even in physics. As Wilhelm Dilthey and Heinrich Rickert noted a century ago, in addition, to the extent that human beings are held to be actors with free will, able to shape history by their intentions, a different sort of indeterminacy is also introduced. This by no means invalidates causal analysis. It merely means that "causation is just one mode of determination, along with self-determination (or spontaneity), chance, and purpose" (Bunge 1996, p. 33).

²³ Somers (p. 775) also says that Kiser and Hechter "create a hybrid standard for explanatory adequacy by conflating the abstract conjunctions of predictive laws with true causality." This seems too strong, though their writing is ambiguous. Rather, I read them as arguing for making use of both predictive, covering-law-style arguments and attempts to identify causal mechanisms in the effort to produce fuller explanations. There seems to be no contradiction in this so long as one keeps them conceptually distinct. Admittedly, Kiser and Hechter are sometimes unclear about how these fit together at the level of philosophical underpinnings. In particular, they do not address the question of what the implications are if (as I think is the case) they see mechanisms working in a probabilistic rather than a strict causation model. This would increase complexity by an order of magnitude and reduce the likely payoff to deductive theorizing—at least until vastly more is known.

CONCLUSION

I have suggested that the methodological divide drawn in this debate is starker than is either widespread in or valuable for the practice of historical sociology. There is room for both general theory and narrative in the explanation of sociologically significant historical phenomena. Moreover, these two do not exhaust useful approaches to such explanation. In particular, though I have not taken the space to develop a positive argument, I have indicated the importance of intrinsically historical theory. This includes both theory that is about the shape of history, its disjunctures as well as its continuities, and theory that is self-conscious about its own historical specificity and that of the phenomena it studies. It is striking that Kiser and Hechter advocate a form of theory that can produce explanations of microprocesses within important historical phenomena, but that will tend to leave their historical importance as such to be accounted for in other ways. Narrative can do some of the work, as Somers suggests, but there is also need for the kind of historical sociology that both learns from and problematizes narrative and that theorizes in nonnarrative forms.

Kiser and Hechter's 1991 article gets us off to a problematic start in thinking about explanation in historical sociology by posing the question in terms of sociology versus history, nomothetic-deductive theory versus inductive or particularizing empirical inquiry. This not only polarizes; it misstates the issues. The fact that Kiser and Hechter are open to arguments based on statistical implication as well as strict causal universality already takes them some distance from the strong ideal of nomothetic-deductive theory as formulated in the *methodenstreit*. By the same token, the terms "idiographic particularism" and "inductivism" poorly describe actual historical research. In the more moderate passages of their contribution published here, Kiser and Hechter indicate that they did not mean to argue one-sidedly that the sort of work they advocated based on general theory was the only good approach to sociological explanation. Their concern, they say, was for the limits of a historical sociology that refused to use such theory. Let us hope they are more worried than they need to be.

One of the drawbacks to thinking in terms of such binary oppositions as general theory versus narrative is that we are not encouraged to examine the questions of whether the strategies may complement each other or whether there are analytic problems that may call for one more than the other. As a familiar example, consider Habermas's (1984, 1988) famous (if problematic) distinction of *lifeworld* from *system*. Part of Habermas's theoretical project was to distinguish dimensions of social life that could be understood well in agent-centered accounts (including, implicitly, nar-

ratives) and those that would be missed or systematically distorted if only understood in such a way. Agent-centered hermeneutic accounts are needed to make sense of friendship; they do a much less adequate job in accounting for the operation of markets (though they are perhaps not irrelevant).

To understand the global expansion of capitalism, for example, is a historical problem. It is not a problem for historians rather than sociologists, because it is basic to the constitution of the contemporary social world and almost everything we study in that world. But it is a problem embedded in time, not just in the absolute but as constituted by the relations among changes in different parts of the system. Some of this will be grasped only by narrative accounts, especially where the experience or intentions of agents are crucial. Global capitalism also includes and indeed is conditioned upon a variety of narrower processes. Some of these can be (and are being) studied fruitfully by means of general theories more or less like those Kiser and Hechter advocate. Others are better addressed through middle range theories, in Merton's (1968) sense. But global capitalism is also importantly a singularity, not merely a cumulation of different instances of smaller, frequently repeated phenomena. It is important for sociology to theorize it as such. This argues for a composite explanatory strategy in which narrative and mechanistic or covering law explanations can each play a part. Nonetheless, an explanation of what capitalism is and how it works will exceed both narrative and discrete causal mechanisms or covering laws.

REFERENCES

- Abbott, Andrew. 1988. "Transcending General Linear Reality" *Sociological Theory* 6:169-86.
- Bhaskar, Roy. 1978 *A Realist Theory of Science*. Hassock: Harvester Press
- . 1986. *Scientific Realism and Human Emancipation*. London: Verso.
- Boudon, Raymond. 1974. *The Logic of Sociological Explanation*. Harmondsworth: Penguin.
- Bunge, Mario. 1996. *Finding Philosophy in Social Science*. New Haven, Conn.: Yale University Press.
- Calhoun, Craig. 1995. *Critical Social Theory: Culture, History, and the Challenge of Difference*. Cambridge, Mass.: Blackwell.
- . 1997a. *Nationalism*. Minneapolis: University of Minnesota Press.
- . 1997b. "The Rise and Domestication of Historical Sociology." Pp. 305-38 in *The Historic Turn in the Human Sciences: Essays on Transformations in the Disciplines*, edited by T. MacDonald. Ann Arbor: University of Michigan Press.
- Coleman, James. 1984. *The Asymmetric Society*. Syracuse, N.Y.: Syracuse University Press.
- Danto, Arthur. 1965 *Analytic Philosophy of History*. Cambridge, Mass.: Harvard University Press.

- Dray, William 1964. *Laws and Explanations in History*, rev. ed. Oxford: Oxford University Press.
- Durkheim, Émile. (1912) 1965. *The Elementary Forms of the Religious Life*. New York: Free Press.
- Elster, Jon. 1985. *Making Sense of Marx*. Cambridge: Cambridge University Press.
- Gallie, W. B. 1968. *Philosophy and the Historical Understanding*, rev. ed. New York: Schocken.
- Goldstone, Jack. 1991. *Revolution and Rebellion in the Early Modern World*. Berkeley and Los Angeles: University of California Press.
- Goldthorpe, John. 1991. "The Uses of History in Sociology: Reflections on Some Recent Tendencies." The 1989 T. H. Huxley Memorial Lecture. *British Journal of Sociology* 42 (2): 211-30.
- Goodwin, Jeffrey. 1997. "The Libidinal Constitution of a High-Risk Social Movement. Affectual Ties and Solidarity in the Huk Rebellion, 1946-54." *American Sociological Review* 62:53-69.
- Goudge, T. A. 1958. "Causal Explanations in Natural History." *British Journal for the Philosophy of Science* 9:194-202.
- Gould, Roger. 1995. *Insurgent Identities: Class, Community, and Protest in Paris from 1848 to the Commune*. Chicago: University of Chicago Press.
- Habermas, Jürgen. 1984, 1988. *Theory of Communicative Action*, 2 vols. Boston: Beacon.
- Hempel, Carl G. 1942. "The Function of General Laws in History," *Journal of Philosophy*, vol. 39.
- . 1962. "Deductive-Nomological vs. Statistical Explanation." In *Scientific Explanation, Space, and Time*, edited by H. Feigl and G. Maxwell. Minnesota Studies in the Philosophy of Science, vol. 3. Minneapolis: University of Minnesota Press.
- Hollis, Martin. 1994. *The Philosophy of Social Science*. Cambridge: Cambridge University Press.
- Hume, David. (1748) 1902. "Enquiry into Human Understanding." In *Enquiries Concerning the Human Understanding and Concerning the Principles of Morals*, edited by L. A. Selby-Bigge. Oxford: Oxford University Press.
- Kiser, Edgar, and Michael Hechter. 1991. "The Role of General Theory in Comparative-Historical Sociology." *American Journal of Sociology* 97 (1): 1-30.
- Kuhn, Thomas. (1962) 1970. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.
- Lakatos, Imre. 1978. *The Methodology of Scientific Research Programs*. Cambridge: Cambridge University Press.
- Marx, Karl. (1867) 1976. *Capital*, vol. 1. Harmondsworth: Penguin.
- Merton, Robert K. 1968. *Social Theory and Social Structure*. New York: Free Press.
- Nagel, Ernest. 1961. *The Structure of Science*. New York: Harcourt, Brace & World.
- Novick, Peter. 1988. *That Noble Dream: The "Objectivity Question" and the American Historical Profession*. Cambridge: Cambridge University Press.
- Pickering, Andrew. 1997. *The Mangle of Practice*. Chicago: University of Chicago Press.
- Popper, Karl R. (1934) 1959. *Logic of Scientific Discovery*. London: Hutchinson.
- . 1972. *Objective Knowledge: An Evolutionary Approach*. Oxford: Oxford University Press.
- Postone, Moishe. 1993. *Time, Labor, and Social Domination*. Cambridge: Cambridge University Press.
- Putnam, Hilary. 1987. *The Many Faces of Realism*. Boston: Open Court.
- Quadagno, Jill, and Stan Knapp. 1992. "Have Historical Sociologists Forsaken Theory? Thoughts on the History/Theory Relationship." *Sociological Methods and Research* 20 (4): 481-507.
- Ricoeur, Paul. 1984. *Time and Narrative*. Chicago: University of Chicago Press.

- Skocpol, Theda. 1979. *States and Social Revolutions*. Cambridge: Cambridge University Press.
- . 1992. *Protecting Soldiers and Mothers: The Political Origins of Social Policy in the United States*. Cambridge, Mass.: Harvard University Press.
- . 1994. "Reflections on Recent Scholarship on Social Revolutions and How to Study Them." Pp. 301–43 in *Social Revolutions in the Modern World*, edited by Theda Skocpol. Cambridge: Cambridge University Press.
- Skocpol, Theda, and Margaret Somers. 1980. "The Uses of Comparative History in Macrosocial Inquiry." *Comparative Studies in Society and History* 22 (2): 174–97.
- Smith, Dennis. 1991. *The Rise of Historical Sociology*. Philadelphia: Temple University Press.
- Stinchcombe, Arthur. 1978. *Theoretical Methods in Social History*. New York: Academic Press.
- Taylor, Charles. (1971) 1985. "Interpretation and the Sciences of Man." *Philosophy and the Human Sciences: Philosophical Papers*, 2, edited by C. Taylor. Cambridge: Cambridge University Press.
- Thompson, E. P. 1968. *The Making of the English Working Class*, rev. ed. Harmondsworth: Penguin.
- . 1978. "The Poverty of Theory or An Orrery of Errors." In *The Poverty of Theory and Other Essays*, edited by E. P. Thompson. New York: Monthly Review Press.
- Tilly, Charles. 1984. *Big Structures, Large Processes, Huge Comparisons*. New York: Russell Sage Foundation.
- . 1995. "To Explain Political Processes." *American Journal of Sociology* 100 (6): 1594–1610.

Commentary and Debate

To conserve space for the publication of original contributions to scholarship, the comments in this section must be limited to brief critiques. They are expected to address specific errors or flaws in articles and reviews published in the *AJS*. Comments on articles are not to exceed 1,500 words, those on reviews 750 words. Longer or less narrowly focused critiques should be submitted as articles. Authors of articles and reviews are invited to reply to comments, keeping their replies to the length of the specific comment. The *AJS* does not publish commenters' rebuttals to authors' replies. We reserve the right to reject inappropriate or excessively minor comments.

RAWLS, DURKHEIM, AND CAUSALITY: A CRITICAL DISCUSSION

Anne Rawls's article, "Durkheim's Epistemology: The Neglected Argument" (*AJS* 102 [1996]: 430–82), challenges social constructivist readings of Durkheim's *Elementary Forms of Religious Life* ([1912] 1995). Most of Durkheim's interpreters have concentrated only on his sociology of knowledge, which concerns the origins of the systems of classification and measurement of various societies. Rawls argues that there is also to be found in this work a much-neglected epistemology, which concerns the warrant, justification, or "empirical validity" of the basic categories of thought, including time, space, classification, force, causality, and totality. She appears to accept that Durkheim has shown that the categories are "valid" because they are "perceived directly as social or moral forces during the enactment of social (religious) practice" (p. 438).¹

Rawls, I will argue, overestimates the success of a causal, perceptual strategy for justifying the categories. It is not even clear that Durkheim in fact attempted to justify the categories in this way. There is implicit in

¹ Rawls (1996, p. 441) uses the term "direct" perception repeatedly and also speaks of social forces being "immediately available to perception."

Durkheim's work a different, potentially more successful epistemological strategy, according to which the categories are justified to the extent that they successfully fulfill their social functions. As Durkheim's theory of the social functions of the categories is part of his sociology of knowledge, his epistemology is thus inseparable from his sociology of knowledge after all. In its emphasis on the universal character of these categories, however, it avoids the conceptual relativist implications of social constructivism that Rawls finds objectionable.

According to Rawls, Durkheim did not attempt to show "that particular social forms created specific conceptual systems" (p. 441). Rather, he argued for "the much stronger claim that there is a direct relation between specific sorts of social practices and particular categories of the understanding" (Rawls 1996, p. 441). In support of her interpretation, she analyzes in detail his accounts of the categories of causality and classification. Causality originates in imitative rites, she explains, and classification in totemic rites. On her reading, causality appears to be the more basic category. The category of causality is rooted in the experience of moral forces, while that of classification derives from the more specific experience of the moral forces associated with the sacred and the profane. Hence, the category of classification depends on that of causality.

Because the category of causality appears to be the more fundamental, I will make it the focus of my discussion. In her account, Rawls appears to be unaware that Durkheim's attempt to ground causality in the experience of moral forces leaves him open to David Hume's objection that one cannot perceive these forces or causes but only their effects. The experience of "moral forces" fares no better than any other kind of internal experience as a justification of the categories. Contrary to what Rawls says (pp. 442–43, n. 9), Durkheim was not well versed in Hume's arguments regarding the origins of our idea of causality. Durkheim had only indirect knowledge of these arguments through philosophy textbooks that provided only incomplete accounts of them. Hence, Durkheim never came to grasp the full import of Hume's argument that forces or causes can be known no more by introspection than by external observation.

HUME ON CAUSALITY

Hume's interest in causality grew out of his larger epistemological and semantic concerns. His goal was to extend the Newtonian, empiricist conception of science to "moral" subjects, including the science of the human mind. The Newtonians eschewed the search for hidden causes behind the phenomena, in accordance with their understanding of Newton's famous dictum, "hypotheses non fingo," with which he had tried to avoid the issue

of the causes of gravitational attraction.¹ The most they thought we could hope to attain is probable knowledge of the laws relating observable antecedents and consequences. In extending this Newtonian model to the science of the mind, Hume ([1748] 1966, pp. 14, 23) regarded the laws of the association of ideas as governing mental phenomena, much as Newton's laws govern external phenomena. In either case, the search for the hidden causes behind these laws was meaningless and in vain.

In defending the Newtonian conception of science, Hume raised both the epistemological argument that a knowledge of causal relations is unattainable and the semantic argument that the concept of causal power as applied to nature or even to the mind is meaningless. The two arguments are not unconnected: if our knowledge of causal relations were grounded in experience, then the concept of causality would have an empirical meaning. While the epistemological argument has received more attention from both Kant and anglophone philosophy, as "Hume's problem of induction," the semantic argument appears to have captured the attention of French philosophers of the 19th century.

For Hume, the notion of a causal relationship seems to entail that of a necessary relationship between cause and effect. However, he found that no ideas are "more obscure and uncertain, than those of *power, force, energy* or *necessary connection*" (Hume 1966, pp. 61–62; Hume's emphasis). Hume regarded these terms as synonymous, apparently using the term "power" not in the sense of a generalized ability to do work but in that of a specific capacity to bring about a certain kind of effect. In that sense, the notion of a necessary relationship to that effect would be implicit in the notion of a power. To clarify the meaning of these ideas, he set out to trace them to their source in either external or internal experience. In both *A Treatise of Human Nature* ([1739] 1975, pp. 155–62, 632–33) and *Enquiries Concerning the Human Understanding and Concerning the Principles of Morals* (1966, pp. 63–69), Hume argued that the idea of power or necessary connection cannot be perceived in any single instance of the action of bodies on each other. Although it is possible to receive kinetic impressions of, for instance, one billiard ball colliding with another, we have no sense impressions that are dynamic. Nor can the idea of power be perceived internally in any single instance of the action of the will upon either the parts of the body or ideas in the mind.

The most relevant to the issue of the introspection of social forces are those of Hume's arguments directed against the hypothesis that the idea of power derives from our internal experience of the action of the will. He maintained that we are able to perceive only the effects of the will

¹ Of course, Newton himself did not always follow this model, having proposed the hypothesis of a gravitational ether in Query 21 of the *Opticks*.

and not the means by which it brings them about. For instance, we perceive only the raising of the arm and not its causes in the action of the mind on the nerves and muscles (Hume 1966, pp. 66–67). Similarly, we are unable to perceive the way in which the mind calls up ideas: “We only feel the event, namely, the existence of an idea consequent to a command of the will; but the manner in which this operation is performed, the power by which it is produced, is entirely beyond our comprehension” (Hume 1966, p. 68).

The issue for Hume is that the contents of our ideas are inactive or inert. Philosophers at least since Descartes had distinguished two senses of the term “idea.” In one sense, an idea is a type of mental activity, and in the other sense an idea is identified with its contents, that is, with an object present to the mind. “Representation” shares the same ambiguity, meaning both the act of representing and the object represented.¹ Hume’s point is that when we introspect, it is only the mental objects and not the mental activity that we see. These mental objects are no more dynamic than those derived from external sensation. As Hume (1975, p. 160) explained, “no impression, either of sensation or reflection, implies any force or efficacy.”

Hume’s argument that the contents of our ideas are purely passive undercuts his own suggestion that the source of this idea of power or necessity is in the feeling of expectation that results from experiencing the constant conjunction of cause and effect (Hume 1966, pp. 74–75; 1975, pp. 163–64). If this feeling of expectation could be dynamic, so could other internal impressions. However, if the objects of our experience are purely passive, as Hume argued, then the notion of a dynamic power must be created by the human mind and not derived from experience. Yet, this would not imply that the concept of causality is meaningless. That conclusion would follow only if we defined causality in terms of power and necessity and adopted a semantics that grounds the meanings of words in experience.

RAWLS’S ACCOUNT OF DURKHEIM ON CAUSALITY

According to Rawls (p. 445), Durkheim solved Hume’s problem of finding an experiential ground for the category of causality by arguing that causality can be “immediately perceived” in single instances of social phenomena: “By substituting enacted social practice and perceptions of moral forces generated in and through those practices by the individual partici-

¹ Other terms that share this ambiguity include *sensation*, *perception*, *conception*, *cognition*, *intuition*, and *notion*.

pants enacting them for individual perceptions of natural phenomena, Durkheim avoids Hume's skeptical conclusion. Enacted practices do display their causes in their effects" (p. 446).

Religious rites are the specific type of social practice in which the participants are supposed to be able to perceive causal powers. According to Rawls, Durkheim believed that participation in rites provides one with "the experience of necessary force" (p. 446). This experience creates feelings of well-being and "moral unity" that can be known by "direct" internal reflection and that are the source of our category of causality. She supposes that our internal experience has some advantage over external experience that allows us to perceive causal forces: "Durkheim argues that, because these feelings are manifest internally, they are immediately available and, therefore, can be known directly rather than indirectly as with external objects. Knowledge of these internal states is, therefore, better and more valid than knowledge of external states or affairs . . . Whereas natural forces cannot be perceived directly, social forces can, and the perception of them can therefore be validly shared with others" (pp. 450–51).

However, if social causes operated internally, Hume's arguments against the view that the concept of causality has its origins in our internal experience would still apply. These arguments would imply that people can introspect only the effects of social causes and not the causes themselves. Rawls (p. 447) in fact concedes this point when she says that the feeling of well-being experienced by the participants in a rite is merely the *result* of that rite. She even quotes Durkheim as having said that the causes of this feeling are not "clearly" seen (p. 449; cf. Durkheim 1995, p. 363). If, however, we were able to experience the causal power of social forces, we would be able to experience other powers as well, and there would be no reason to appeal to specifically social forces to ground the concept of causality.

There is a second way in which Rawls equivocates as well. She sometimes says such things as "The causal *relation* is a social *force* available in direct perception" and "immediately displayed in its effects" (p. 449; emphasis added). Unlike Hume, who regarded the notions of causal force and causal relation as synonymous, Durkheim distinguished them. He regarded the notion of force as only a part of the notion of a causal relation, which also includes a notion of necessity distinct from that of force or power (Durkheim 1995, p. 370). Durkheim (1995, p. 367–72) traced only the notion of power or force, not that of necessary relation, to the internal experience of social forces. He derived the concept of necessity from the moral obligation to participate in imitative religious rites. Society imposes this obligation on its members because a social interest is at stake. To obligate someone to imitate an animal or plant so that it will reproduce

and flourish is to presume that performing the rite necessarily leads to this result. If society allowed people to doubt this causal relationship, Durkheim (1995, pp. 370–71) argued, it could not compel them to perform the rite. For Durkheim, then, the notion of force or power has its origins in the simple fact *that* society imposes obligations on us while the notion of a necessary relation has its origins in the nature of this obligation itself. Elsewhere, he assimilates the notion of moral necessity to that logical necessity (Durkheim 1995, p. 17, n. 20). An implication of his sociology of knowledge is that there is no concept of necessity other than that of logical and moral necessity and that what we take to be causal necessity is merely this concept of logical necessity applied to nature.

Finally, Rawls does not seem to recognize that epistemology has to do with the justification of beliefs rather than the origin of our concepts. Hume's investigation into the origin of our idea of causality is a semantic, not an epistemological, inquiry. The justification of a belief has to do with its logical agreement with the facts of experience. A belief can be supported by experience without having been derived from it. Although Durkheim may not have separated semantics from epistemology, he recognized the distinction between questions of origin and questions of justification. He made it clear that he was presenting only an account of the origins of the category of causality and not an epistemological justification of it:

Doubtless, certain philosophers refuse any objective value to this conception; they see in it only an arbitrary construction of the imagination that corresponds to nothing in things. But we do not have to ask ourselves for an instant whether or not it is founded in reality: it suffices for us to verify that it exists, that it constitutes and that it has always constituted an element of the common mentality; and that is what is recognized even by those who criticize it. Our immediate goal is to seek not what it may be worth logically, but how it is to be explained. (Durkheim 1912, p. 519; 1995, p. 367)⁴

Far from Durkheim having provided an epistemological justification of the category of causality in terms of the experience of social forces, there are problems with this even as an account of the origin of this category. Whatever the merits of his account of the notion of necessary relation, it is more of a functional than a causal account, as it appeals to the purpose of imposing obligations on the members of society. In his account of the notion of causal power, Durkheim appears to have been unaware that Hume's arguments would imply that we could not perceive social causes at all, not even obscurely. Durkheim's knowledge of Hume was

⁴ I have provided my own translation of all *Elementary Forms* quotations from Durkheim's original 1912 French text. Reference to the complete 1995 translation is provided for purposes of comparison.

in fact rather incomplete. Rawls is just wrong when she says that "Durkheim uses Hume's arguments with great sophistication." (pp. 442–43, n. 9). In the same footnote, where she argues that Durkheim "would have been familiar with Hume," she generalizes from what may have been true of German university curricula in philosophy in the 19th century to a speculative conclusion about the French curriculum in this period and thus about Durkheim's knowledge of Hume (1996, pp. 442–43, n. 9). Durkheim, however, never made any explicit reference to any of Hume's works. Also, there are no library records at the Ecole Normale Supérieure showing that Durkheim had ever borrowed any of Hume's books (Schmaus 1995, p. 17). Although Rawls is correct to say that "Durkheim intended his epistemology to make traditional philosophy obsolete" (p. 452, n. 14), she has surprisingly little to say about exactly what this tradition was for Durkheim.

THE FRENCH TRADITION

During the Third Republic, French philosophy students had only indirect knowledge of Hume's works through various philosophy textbooks, such as Paul Janet's (1879) *Traité élémentaire de philosophie à l'usage des classes* and Elie Rabier's ([1884] 1893) *Leçons de philosophie* (cf. Brooks 1996, p. 386). Durkheim may have used these texts while teaching philosophy at lycées in the period before 1887. In *The Division of Labor in Society*, which he began while a lycée professor, he cited both Rabier ([1902] 1984, p. 199, n. 5) and Janet [1933] 1964, p. 411). Janet, who was on his dissertation committee (Lukes 1973, p. 297), is cited again in his lectures on the family (Durkheim [1888] 1978, p. 222), and Rabier is also cited in "Individual and Collective Representations" (Durkheim [1953] 1974, p. 5).

These textbooks taught the official spiritualist philosophy that was designed to inculcate patriotic and religious values among the educated elite in France. The philosophical curriculum, which was originally set up by Victor Cousin during the July Monarchy, had four major divisions—psychology, logic, moral philosophy, and metaphysics—which were to be studied in this order. This order is followed not only in the texts cited above but in the recently discovered notes taken by André Lalande as a student in Durkheim's philosophy course at the Lycée de Sens during the academic year 1883–84 (Durkheim 1884).

The philosophical psychology that made up the first part of the French curriculum included an account of the categories and was supposed to provide a foundation for the other branches of philosophy. It drew on the work of Maine de Biran as well as Cousin. Maine de Biran had defended against Hume's criticisms the thesis that the idea of causality derives from

our inner experience of the power of the will over the body. However, he did not reply to Hume's arguments that the idea of causality does not come from our experience of the operation of the will on its ideas, as he did not try to derive the idea of causality from this source. French academic philosophers of this period seemed to find Maine de Biran persuasive. For example, although Rabier ([1884] 1893, pp. 289–92) approvingly quoted long passages from Hume's arguments that the idea of causality does not derive from external experience, he did not give the same consideration to Hume's arguments regarding internal experience but simply accepted Maine de Biran's replies:

Without here going into the examination of the objections that Hume made in advance of this theory, and in granting to Maine de Biran that this experience would in effect be sufficient for engendering the idea of cause, in case we should not already have it, one may say that, before this experience, we already have this idea. In effect, as soon as I have willed the motion, before all commencement of its execution, and even when no execution of the motion takes place, the will appears to me as a possible cause of this effect still to come. (Rabier 1893, p. 295; my translation)

Janet provided a more detailed and fairly accurate account of Maine de Biran's replies to Hume. According to Janet and Maine de Biran, Hume had argued that the influence of the will on the parts of the body is known merely in the same way as are other operations of external nature, that is, as a succession of phenomena. In reply, Maine de Biran distinguished internal from external perception. He argued that although we may doubt the existence of the external objects that cause our perceptions, in internal perception, the very act of perception is its own object. The thing and its appearance are one and the same. The internal perception of activity or power could then serve as a model for the idea of an external force (Janet 1879, p. 201; cf. Maine de Biran 1841, 4:276–77; or 1982, 11:367–68). This first argument appears to beg the question in assuming that our mental activity can be perceived as the object of a representation. At best, introspection could reveal *that* we willed to move a part of the body but could not give us any idea of the will as causally active. Indeed, in a different, posthumously published version of his replies to Hume, Maine de Biran himself denied that we represent the activity of the will and argued that only external perception involves the representation of objects. Thus, even for Maine de Biran, we can have knowledge only of the relation between the will and its effects and not of the will's causal powers (Maine de Biran 1982, 8:229; cf. Moore 1970, pp. 64–67, 85; Morin 1974, pp. 447–48). However, this is not the version of his argument discussed in Janet's textbook and that Durkheim would have known.

Maine de Biran's other arguments also support the view that the introspection of willed effort is the basis for only our idea of causal relations

and not that of causal powers. For instance, he argued that unlike in external observation, in internal observation, repetition adds nothing to our belief in the connection between cause and effect (Janet 1879, p. 201; cf. Maine de Biran 1841, 4:277; or 1982, 11:368). He also argued that we have an "intimate sense" of this connection even if we do not know the means by which the mind controls the body (Maine de Biran 1841, 4:280; or 1982, 11:369–70).

Cousin (1860) and subsequent French philosophers used Maine de Biran's arguments concerning the introspection of the will as a basis from which to defend a traditional, Cartesian notion of a substantial soul. Although, early in his career, Durkheim (1884, pp. 113–14) defended the spiritualist notion that the self or the will is known by introspection, by the time he wrote the *Elementary Forms*, he held that these ideas are social in origin. Far from deriving the category of causality from the experience of the power of the will over the body, he said that the individual's idea of this power derives from the notion of the soul, which originated in the totemic principle and thus has social causes. In a passage clearly directed against arguments like Maine de Biran's, Durkheim wrote:

Thus man could not have arrived at conceiving himself as a governing force of the body in which he resides except under the condition of introducing, in the idea that he makes of himself, concepts borrowed from social life. . . . In fact, it is indeed under the form of the soul that he has always represented the force that he believes himself to be. But we know that the soul is something entirely other than the name given to the abstract faculty of moving, thinking, or sensing; it is, above all, a religious principle, a particular aspect of the collective force. After all, man is conscious of a soul and, consequently, a force because he is a social being. (Durkheim 1912, p. 523; 1995, p. 370)

The spiritualist notion that the action of the will served as the model for the idea of external force was the basic premise of the animist theory of the origin of religious belief, which the *Elementary Forms* goes to great lengths to refute. The animist theory had been defended by Rabier: "Primatively, all the external causes that we conceived were conceived by us on the model of this internal activity, desire, or will, that we grasp within ourselves: thus all of nature is animated; we see life everywhere, everywhere efforts, tendencies, forces; and these invisible forces are for us the secret spring that unfolds all the visible phenomena of nature" (Rabier 1893, pp. 296–97). Providing an alternative, sociological account of the categories was a way for Durkheim to defeat both the animist theory and the spiritualist philosophy with which it was allied.

Nevertheless, although Durkheim broke with the spiritualist tradition, he appears to have carried over to his sociology its assumption that we can introspect causal powers and not merely their effects or causal relations. He presented Hume's arguments that the source of the idea of cau-

sation or power cannot be found in external experience (Durkheim 1995, p. 368) but did not give Hume's arguments that it could not be found in our internal experience of the operation of the will. Instead, Durkheim (1995, pp. 368–69) invented his own arguments, citing disanalogies between the will and causal powers or forces. In these arguments, he was attacking the spiritualist philosophy of Maine de Biran, Cousin, Janet, and Rabier, not the empiricist philosophy of Hume, of which he had no direct knowledge. Hence, Durkheim did not take into account Hume's point that even our ideas of internal reflection are inert, which, as we have seen, has direct implications for Durkheim's theory of the origin of the idea of causal power. Having no doubt studied the same texts, Durkheim's French contemporaries were probably equally unaware of Hume's arguments and did not raise this objection.

CLASSIFICATION

Rawls's argument regarding the category of classification is dependent upon that regarding the category of causality. She says that the category of classification is not a collective representation abstracted from or constructed on the basis of various particular classifications. Instead, it is produced by feelings of moral forces associated with the distinction between the sacred and the profane that are experienced during totemic rites. It is these experiences that ground the validity of the category of classification (Rawls, pp. 453–57). Hence, it would seem that the argument regarding classification is subject to the same objections as is the argument regarding causality: that is, that one cannot perceive moral forces but only their effects.

Also, one has to ask how the experience of moral forces could sometimes give rise to one category and sometimes to another. One might suppose that in one case one perceives these forces, or the context in which they arise, as "different" than in another. However, Rawls (pp. 453, 458) seems to think that for Durkheim, our natural ability to recognize resemblances and differences is not "valid," which would imply that it cannot give rise to distinct, valid categories. It would not suffice for her to reply that perceptual knowledge is "nevertheless in touch with reality," for she clearly says that this being in touch gets us no further than Hume's custom and does not produce valid knowledge (p. 458). The shift to "witnessable" practices (e.g. Rawls, p. 468) will not help her case either, because something cannot be witnessed without being observed.

REMARKS ABOUT THE CATEGORIES IN GENERAL

Durkheim's (1995, p. 13–17) defense of his sociological theory of the categories consists in arguing that it provides a better explanation than any

philosophical theory of their generality, universality, necessity, and social variability. Neither the variability nor the generality of the categories, which has to do with the scope of things to which they apply, presents a problem for traditional empiricism. However, empiricism cannot explain their universality or necessity. By the universality of the categories, he meant that they are independent of individual subjects and serve as "the common ground where all minds meet" (Durkheim 1995, p. 13). Categories, like concepts generally, are "communicable to a plurality of minds, and even, in principle, to all minds" (Durkheim 1995, p. 435, n. 9). He explained that the necessity of the categories is not a physical or metaphysical necessity, since the categories vary with time and place, but a logical necessity, analogous to moral necessity, from which it originated. For Durkheim (1995, p. 17), the necessity with which the categories impose themselves upon our thought is to be explained in terms of society's need for conformity of thought among its members. Again, however, as we saw with regard to his account of causal necessity, this is a functional rather than a causal explanation.

Rawls's "socioempiricism" can no more account for the universality and necessity of the categories than traditional empiricism can. If the categories are in any way grounded in perception or some other cause, even the inner experience of social forces, they are merely contingent, not necessary. That is, this argument would show that they are not even necessary in Durkheim's physical or metaphysical sense, let alone logically necessary. Furthermore, if the categories are contingent upon social causes, and if these causes varied, the categories would be only particular to a certain type of society, not universal. Rawls (p. 451) also appears to have neglected what Durkheim said about the universality of the categories when she attributed to him the belief that "socially derived categories are more valid than the categories of natural science and logic." Far from making this distinction, Durkheim (1913, pp. 35–36) criticized Lévy-Bruhl for doing so, arguing that the categories presupposed in modern scientific thought evolved from primitive religious categories, not in opposition to them.

Admittedly, Durkheim appears to have made conflicting demands upon a theory of the categories when he required that it explain both their universality and their variability. To resolve this conflict, we need to distinguish the collective representations of the categories, which may be socially variable, from the categories themselves, which are universal. Indeed, Durkheim drew something like this distinction in his early philosophy lectures, in his attempt to reconcile Cousin's philosophy to that of Maine de Biran. As he explained the conflict, whereas for Maine de Biran the principle of causality is generalized from inner experience, for Cousin, the principle is *a priori*, but the "idea" of causality is given experientially. Cousin's account then raises the problem for Durkheim: how can the prin-

ciple be a priori if the ideas contained in it are not? To resolve this difficulty, Durkheim distinguished these ideas as given by reason from the same ideas given by experience. For instance, he said, reason compels us to relate phenomena to something else but does not say what that something is. Experience intervenes and provides the concrete representation of the idea of substance. Similarly, reason provides only the idea of a necessary antecedent of a phenomenon. That which is exactly a cause, he argued, only internal experience can show us (Durkheim 1884, pp. 138–39). In sum, according to Durkheim, reason provides us with purely abstract syntactic categories such as subject and antecedent. Experience alone—whether inner or outer—allows us to represent these categories in a concrete way as substance and cause. Although he did not yet have the concept of a collective representation when he gave these lectures, this concept could be assimilated to his early notion of ideas of experience.

As we saw above, Rawls distinguished the category of classification from its collective representation. It would clarify her position if she made a similar distinction in the case of the other categories as well. She says that the categories are “immediately available” to perception by being “inherent” in the rites themselves (pp. 445–46). But what could “inherent” mean in this context? If the category of causality were already present in the rite, how could it be said to have been caused by the rite or by the experience of the effects of the rite? At best, the rite could give rise only to the collective representation of the category of causality, not to the category of causality itself. Also, if the categories are supposed to be grounded in the experience of moral forces, it is difficult to see how she can distinguish the categories from some sort of representation, idea, or mental content, whether individual or collective, as she appears to want to do in her implicit criticism of William Dennes (1996, p. 464).

To distinguish the categories from their collective representations is not necessarily to leave the categories themselves as some sort of free-floating entity. These concepts may be defined by their social functions. For example, the concept of causality may serve the function of making it possible to ascribe blame and responsibility, thus allowing for a society to develop moral rules. If one insists that these functionally defined concepts be represented somewhere, one could think of them as concepts developed by a community of social scientists and philosophers for the purpose of interpreting the collective representations of other cultures.

A FUNCTIONAL ACCOUNT OF THE CATEGORIES

Implicit in the *Elementary Forms* is a functional account of the categories. Durkheim (1995, p. 18, n. 24) compared the categories to tools that have been improved over time. He also argued that society is not possible with-

out the categories. For there to be a society, individuals must be organized into groups. Society thus requires the category of a class. Places must be divided among these groups, requiring divisions and directions of space. Convocations to feasts, hunts, and military expeditions require a common way of fixing dates and times. Cooperation with the same end in view requires agreement about means and end, that is, agreement about a causal relationship (Durkheim 1995, pp. 444–45).

Perhaps Durkheim intended a functional account of the categories to provide the basis for their epistemological justification. The relation between specific social practices and particular categories could be conceived as a functional one. Rather than seeing the categories as caused by religious rites, we could see them as having the function of making such rites, and thus the social feelings they generate, possible. These social feelings “feed back” to help preserve the society and thus the society’s collective representations of the categories. These collective representations of the categories would be then justified to the extent that they fulfill these functions.

Indeed, it is only the collective representations of the categories and not the categories themselves that are susceptible of epistemological justification. Only beliefs or propositions in which two or more concepts are connected can affirm or deny something and hence are capable of being justified. A category is an individual concept that affirms nothing. However, we can think of the collective representations of the categories as making implicit affirmations. For instance, the collective representation of space as three dimensional affirms that it is such. Also, although the category of genus by itself affirms nothing, the collective representation of this category in a classification system entails many affirmations.

A functional justification of a society’s collective representations of the categories would differ from a traditional epistemological justification in being broadly pragmatic rather than foundationalist. Indeed, it is not at all clear that Durkheim intended a foundationalist justification, given what he said about our collective representations of the categories being like tools improved over time. If they were already well-founded, why would they need to be improved? Although Durkheim criticized the pragmatists, he shared their abandonment of a foundationalist view of knowledge. As I have argued elsewhere, he subscribed to the view that scientific knowledge is irreducibly conjectural (Schmaus 1994, pp. 89–110). He had in fact defended the use of the hypothetico-deductive method, even in philosophy, as early as his *Sens* lectures (Durkheim 1884, pp. 13–14, 19).

Rawls, however, appears to be seeking a foundation for the categories in their social causes and thus setting herself to the nearly insurmountable task of succeeding where no previous epistemology ever has. The functional justification I am proposing, on the other hand, avoids many of the

philosophical problems encountered by Rawls's causal account. It escapes the difficulty of not being able to perceive causes acting. Nor does it muddy the distinction between semantics and epistemology. Also, unlike Rawls's socioempiricism, my approach can explain the necessity and universality of the categories. Rather than being the contingent effects of social causes, they are the necessary conditions for social life, and hence they are to be found in all societies. Finally, Rawls's causal account seems to found the concept of causality upon a mistake or a confusion: The imitative rite does not produce its intended effect. Rather, the participants think it is working because the rite makes them feel good. It is no objection to a functional account of some social fact, however, that the beneficial consequences to which it appeals are unintended.

Indeed, some of Rawls's own language is functionalist, as for instance when she says that the categories' "purpose is to make intelligibility and social cooperation possible" (p. 452) or that "the feeling of moral unity is the purpose of the ritual" (p. 451). Her account of the evolution of the categories also uses functionalist language: "The categories did not evolve to fulfill philosophical purposes. The categories and the social practices that produce them came into being to fulfill the social need for shared categories of the understanding among members of the same group" (p. 452). She never actually showed that social practices "produce" the categories. In fact, as we saw above, they presuppose them. What these practices produce, at best, are collective representations that fulfill the functions of the categories. When she says the categories "came into being to fulfill the social need," this can mean only that certain collective representations were selected because they served these categorial functions.

Rawls (p. 439) regards as a part of Durkheim's *epistemology*, rather than his sociology of knowledge, his account of how the categories make communication possible. Hence, it would appear to be consistent with her own views to consider as relevant to their justification the issue of how well collective representations fulfill their functions. However, a theory of the functions of the categories would appear to belong to the sociology of knowledge as well. Durkheim's epistemology would thus appear to be inseparable from his sociology of knowledge.

Durkheim's epistemology provides a "naturalized" theory of the categories in the sense of an empirically defeasible account. It can be tested by determining whether in fact every society does have collective representations of the categories of causality, classification, space, time, and so on. If we find a society that lacks one of these categories, that would disconfirm Durkheim's hypothesis that this category is universal and necessary for social life. Indeed, the application of empirical methods to epistemological questions is at least part of what he meant when he said that his sociology would renew philosophy.

REFERENCES

- Brooks, John L, III. 1996. "The Definition of Sociology and the Sociology of Definition: Durkheim's *Rules of Sociological Method* and High School Philosophy in France." *Journal of the History of the Behavioral Sciences* 32 (4): 379-407.
- Cousin, Victor. 1860. *Du Vrai, du beau, et du bien*. Paris: Didier et Cie.
- Durkheim, Émile. 1884. *Cours de philosophie fait au Lycée de Sens*. Bibliothèque de la Sorbonne, manuscript no. 2351. Also available from the University of Wisconsin, Madison, Microforms Center, film no. 9307.
- . 1912. *Les Formes élémentaires de la vie religieuse*. Paris: Quadrige/Presses Universitaires de France.
- . 1913. Combined review of "Lévy-Bruhl, L., *Les Fonctions mentales dans les sociétés inférieures*," and "Durkheim, Émile, *Les Formes élémentaires de la vie religieuse*." *L'Année sociologique* 12:33-37.
- . (1933) 1964. *The Division of Labor in Society*, translated by George Simpson. New York: Macmillan.
- . (1953) 1974. "Individual and Collective Representations." Pp. 1-34 in *Sociology and Philosophy*, translated by D. F. Pocock. New York: Free Press.
- . (1888) 1978. "Introduction to the Sociology of the Family." Pp. 205-28 in *Émile Durkheim on Institutional Analysis*, edited and translated by Mark Traugott. Chicago: University of Chicago Press.
- . (1902) 1984. *The Division of Labor in Society*, translated by W. D. Halls. New York: Free Press.
- . (1912) 1995. *The Elementary Forms of Religious Life*, translated by Karen Fields. New York: Free Press.
- Hume, David. (1748) 1966. *Enquiries Concerning the Human Understanding and Concerning the Principles of Morals*. Oxford: Clarendon Press.
- . (1739) 1975. *A Treatise of Human Nature*. Oxford: Clarendon Press.
- Janet, Paul. 1879. *Traité élémentaire de philosophie à l'usage des classes*, 9th ed. Paris: Librairie Ch. Delagrave.
- Lukes, Steven. 1973. *Émile Durkheim: His Life and Work*. New York: Penguin Books.
- Maine de Biran, Pierre. 1841. *Oeuvres philosophiques de Maine de Biran*, 4 vols. Edited by Victor Cousin. Paris: Ladrangé.
- . (1920-49) 1982. *Oeuvres complètes*, 14 vols. Edited by Pierre Tisserand. Geneva: Slatkine.
- Moore, Francis Charles Timothy. 1970. *The Psychology of Maine de Biran*. Oxford: Clarendon Press.
- Morin, Serge J. 1974. "Maine de Biran: A New Dualism." *The Philosophical Forum* 5:441-59.
- Rabier, Elie. (1884) 1893. *Leçons de Philosophie: I. Psychologie*, 4th ed. Paris: Librairie Hachette et Cie.
- Schmaus, Warren. 1994. *Durkheim's Philosophy of Science and the Sociology of Knowledge: Creating an Intellectual Niche*. Chicago: University of Chicago Press.
- . 1995. "Rabier as Durkheim's Source for Hume's Arguments." *Durkheimian Studies/Études durkheimiennes* 1:17-23.

WARREN SCHMAUS

Illinois Institute of Technology

DURKHEIM'S CHALLENGE TO PHILOSOPHY: HUMAN REASON
EXPLAINED AS A PRODUCT OF ENACTED SOCIAL PRACTICE

In my article, "Durkheim's Epistemology: The Neglected Argument," I argue that Durkheim articulated a sociologically based epistemology intended to replace philosophy as the arbiter of claims to knowledge, truth, and validity. In "Rawls, Durkheim, and Causality: A Critical Discussion," Warren Schmaus criticizes my argument, claiming instead that Durkheim articulated a sociology of knowledge based on "representations," not an epistemology in the classical sense. While acknowledging Durkheim's claim to have solved Hume's empiricist epistemological dilemma, Schmaus states that Durkheim's position amounts to only a "justification from functions," within a conventional epistemological context.

In making his criticisms, Schmaus dismisses my argument in its entirety as the result of ignorance, on both my part and on the part of Durkheim, of what is involved in an epistemological argument. This mirrors the treatment Durkheim experienced when he published the original argument in 1912. Scholars have consistently refused to entertain seriously the idea that Durkheim intended to articulate a socially based epistemology in the classical sense. The philosophical idea that the social is necessarily contingent is treated as dogma. Schmaus takes the accepted view that Durkheim simply did not make an epistemological argument.

It has been my aim to give serious consideration to the possibility that Durkheim did, in fact, articulate a socially based epistemology. Rather than accepting the traditional view that positing a social basis for reason necessarily resulted in contingency, Durkheim argued that requirements of reason and intelligibility placed demands on social practices that were constant enough across societies to explain the development of more or less the same faculties of reason in all societies. The implications of such an argument would be revolutionary. Treating the development of reason as a *result* of social processes, rather than as a *cause*, would transform social processes, generally thought to introduce an inescapable element of contingency, into the origin of a more or less universal reason. The argument places sociology at the heart of essential philosophical debates.

Attempts to establish a foundation for epistemology within sociology are evident not only in Durkheim's argument but in the work of other classical social theorists as well. In my view, the contemporary dissatisfaction with "theory" proper and the turn to postmodern, poststructural, and sociology-of-knowledge arguments is due to the neglect and misinterpretation of earlier attempts to articulate a sociologically based epistemology, leaving sociologists unnecessarily at the mercy of the philosophical assumption that "social" equals "contingent."

Schmaus's critique recapitulates these historical misunderstandings. As far as Schmaus is concerned, the central argument of Durkheim's magnum opus was never written: Durkheim remaining to the end a rationalist in the classical sense, who did not criticize the inherent universal "givenness" of reason. As a rationalist, Durkheim supposedly remained ignorant of fundamental empiricist arguments, considered reason part of the given order of things, and never made an argument for the empirical origins of reason. In Schmaus's view, Durkheim should be interpreted as arguing that social practices create particular contents for preexisting universal categories, not as making, what Schmaus calls, a "foundationalist" epistemological argument for the origin of the categories of the understanding themselves.

This interpretation, he claims, keeps the social contribution to the categories, which Schmaus believes is necessarily contingent, to a minimum, retaining thereby an idea of universality and necessity that he says my interpretation lacks. Schmaus's assumption that a social origin for the categories necessarily renders them contingent speaks to the crux of the matter. From within a traditional epistemological context, the social contribution to reason is necessarily contingent and must be kept to a minimum. This is why, prior to Durkheim, foundationalist arguments never worked. From within a classical philosophical context, social practices must be treated as merely derivative in importance, and epistemological argument is generally built from individual perception.

However, and this is critical to understanding the coherence of sociology as a discipline, Durkheim directly challenged that classical assumption. Durkheim was *not* talking about social contents filling universal rational containers. He did not begin with individual reason but, rather, with those social beings that are created through the enacted practices of assembled groups. Durkheim argued that the social created the container, the framework, the category itself, not only the contents. This does not make Durkheim a forerunner of postmodern relativism. He did *not* accept the philosophical premise that a social origin for reason rendered it relative. He argued that there were inherently universal elements involved in the social creation of reason. His argument was revolutionary. It was meant to challenge philosophy's claim to define epistemology.

In addressing Schmaus's criticisms, I have separated the issues into seven main points. *First*, Schmaus claims that foundational arguments regarding the origin of categories, like the one I attribute to Durkheim, have nothing to do with epistemology, which he identifies with justified belief; *second*, Hume, he says, made both a semantic and an epistemological argument. He treats the argument that general ideas can only be inferred as the epistemological argument, arguing that Hume's search for the origins of the concept of causality constitutes only a semantic argu-

ment, the implication being that Durkheim's argument with regard to causality, because it follows Hume's in form, is not epistemological either; *third*, he argues that Durkheim did not make a classical or "foundational" epistemological argument for the origin of concepts. Instead, he says, Durkheim assumed an underlying rationalism and then made an argument about the social functions of knowledge that did not make claims to either the origin or the validity of knowledge; *fourth*, Schmaus claims that Durkheim's argument that "feelings" produced by social practices, available as internal experiences of the moral force of those events, have a special status as knowledge, is an argument from introspection of the sort that Hume's argument against introspection refutes; *fifth*, he says that Durkheim was unaware that his position was refuted by Hume because he was ignorant of Hume's arguments. Durkheim, he says, was only familiar with classical French philosophy, and his arguments can only be understood in that context; *sixth*, Schmaus states that there is no difference between social and natural causes. Since natural causes cannot be directly perceived, he says it follows that social causes cannot be directly perceived; and, *seventh*, he argues that all collective representations, including the six Durkheim calls categories of the understanding, are the particular contingent content of universal categories. Durkheim is better understood, according to Schmaus, as having provided a functional account of the content of these categories and the way in which societies produce their contingent manifestations through external social constraint.

EPISTEMOLOGY: ARGUMENT FROM ORIGINS OR JUSTIFIED BELIEF?

In my article, I argued that Durkheim's epistemology consists of the argument that the categories of the understanding—which, as elaborated in *The Elementary Forms* (Durkheim [1912] 1915), are six in number and the same six in every society—have their origin in the experience of participation in the enactment of social practices, not in individual perceptions of natural phenomena. Schmaus seems to think that in claiming this I make the rather enormous and naive mistake of misunderstanding what an epistemological argument is altogether. Schmaus writes, "Finally, Rawls does not seem to recognize that epistemology has to do with the justification of beliefs rather than the origin of our concepts" (p. 851).

The problem with Schmaus's argument in this regard is that the justified belief position is a derivative of the classical epistemological debate. For most of the history of philosophy, epistemological arguments proceeded by trying to determine the origin of concepts. These arguments were not semantic but, rather, sought to demonstrate the valid *causes* of

ideas. Early philosophers often worked with a model of the individual confronting an external reality and a picture, or copy, theory of the origin of ideas. They did not believe that the boundaries of knowledge were identical with a universe of contingent concepts and the relationship between those concepts, as in a language. Empiricist philosophers argued that ideas were valid only if they could be shown to be accurate copies, or impressions, of something external that had real existence.

In contemporary philosophy, a debate has arisen over the meaningfulness of this classical type of argument from origins. Since Frege and Russell at the turn of the century, and particularly following Wittgenstein's argument that words do not refer to "things" or to ideas of things but get their meaning from the conventions of language games, the argument has increasingly been that the limits to knowledge are semantic. Many philosophers have urged a quasi-pragmatic view of the semantic limits of knowledge that reduces epistemology to a concern with justified belief. The contemporary semantic project has sometimes been referred to as the "linguistic turn" in philosophy. On this view, to understand a sentence is to know its truth conditions, and these can be entirely customary.

While epistemology has for this reason taken something of a back seat in contemporary philosophy, for most of the history of philosophy, epistemological questions were a primary focus of disciplinary argument. There was a time when philosophy was almost synonymous with epistemology in the classical sense. Schmaus confuses this contemporary emphasis on justified belief for what counted as epistemology throughout the history of philosophy. The contemporary debate over whether classical epistemological arguments like Hume's were really justified belief arguments is essentially irrelevant to the matter at hand. What matters is how Durkheim would have interpreted Hume. In the 18th and 19th centuries, Hume was interpreted as having articulated a classical epistemological argument that "proved" that arguments from origins fail. It was this interpretation that led to the skepticism about epistemology that ushered in the new arguments for justified belief. It is precisely because his arguments were *interpreted* as foundational in the 19th century that Hume marks an important turning point in the study of epistemology.

IS HUME'S ARGUMENT FROM INFERENCE DIFFERENT FROM HIS SO-CALLED SEMANTIC ARGUMENT?

According to Schmaus, Hume made two arguments, one semantic and one epistemological. When Schmaus refers to Hume's epistemological argument, he means the argument that general ideas cannot be derived directly from experience, only inferred. When he refers to Hume's semantic

argument, he apparently means to single out the argument for the origin of the concept of "causality." Schmaus hopes to establish that if Durkheim is understood as following Hume, then he must also be making a semantic argument and not an epistemological argument from origins. The discussion of the origin of the concept of causality is semantic, according to Schmaus, because it is a search for the "meaning" of a word. The problem with this way of interpreting Hume is that his so-called semantic argument for the origin of the concept of causality is the mainstay of his epistemological argument that general ideas can only have their origin in inference. Discussions of the argument from inference, both in the original and in commentaries, invariably refer to the discussion of the origins of causality as the "proof" that general ideas can only be inferred.

Furthermore, Hume does not treat the origin of the concept of causality as a matter of semantics. He considers establishing the "meaning" of the concept of causality to be a matter of establishing its origins, "we must consider the idea of *causation* and see from what *origin* it is deriv'd" (Hume [1739] 1975 pp. 74–75; emphasis added). For Hume, as for Locke, the causal origins of ideas were their "true meaning." Hume writes, "Tis impossible perfectly to understand any idea, without tracing it up to its *origin*, and examining that primary *impression*, from which it arises" (Hume 1975, pp. 74–75; emphasis added). The origins of ideas are, for Hume, to be sought in sense impressions of external reality, not in relations between words or propositions, as Schmaus suggests. Semantic theories look for the meaning of words in the conventions of grammar, language games, or other framing conditions for the use of propositions, whereas, in this case, Hume looked for the cause that created the mental impression corresponding to the word.

IS DURKHEIM'S ARGUMENT EPISTEMOLOGICAL IN THE CLASSICAL SENSE OF THE WORD?

I referred to Durkheim's argument as an epistemology because he attempted to provide for the empirical validity of a set of ideas common to persons by discovering their origins in experience. Schmaus claims that Durkheim does not make an epistemological argument from origins at all but, rather, a sociology of knowledge argument from "functional" considerations. Durkheim certainly did articulate a sociology of knowledge. That is not a point of dispute. What we disagree over is whether Durkheim also articulated an epistemological position that can be distinguished from his sociology of knowledge.

I believe that Schmaus is able to deny the existence of a separate epistemological argument only by attributing to Durkheim an underlying a priori rationalism. In different parts of his text, Schmaus interprets Durk-

heim's reference to the categories as "the framework of the intelligence" to mean *both* that Durkheim believes that the concepts that become the framework and the limit of knowledge are related to societies in a historical and contingent fashion and that Durkheim assumes an inherent rationalist framework. Schmaus cites Durkheim's early work as evidence that he took an a priori rationalist position. But, Durkheim makes it clear that whatever his views on rationality may have been early on, they changed dramatically after the writings of 1884 cited by Schmaus. Durkheim ([1913-14] 1960, p. 71) writes that "we can no longer accept a single, invariable system of categories or intellectual frameworks." According to Durkheim (1960, p. 67), "all that constitutes reason, its principles and categories, has been made in the course of history." He faults Kantian philosophers, whom he calls apriorists, for not concerning themselves sufficiently with origins, saying the apriorist position, which Schmaus attributes to him, is a "lazy man's solution" (Durkheim 1915, p. 172), providing neither "explanation or justification" for the categories (Durkheim 1915, p. 27).

Neither the sociology of knowledge or the a priori rationalist interpretation reflects Durkheim's argument with regard to those six ideas that he elaborates as categories of the understanding. These categories have empirical validity, they are not historically contingent, they are not a priori, and a society cannot exist unless it creates them. Durkheim (1915, p. 22; emphasis added) argued that the study of primitive religious beliefs would reveal the origin of the same categories we have today: "Now when primitive religious beliefs are systematically analyzed, the principal categories are naturally found. They *are born in religion* and of religion; they are a *product of religious thought*." Because they are necessary, they have origins in social processes that are also necessary and must occur in essentially the same way in all societies. Therefore, the origin of each of the categories will be in essential respects the same in every society. Both in the text and in a footnote, Durkheim discusses what he means by origins and contrasts his view with historical contingency. He is careful to say that his is not a search for a historically first or specific and unique point of origin, but "what we want to do is to find a means of discerning the *ever-present causes*" of religious practice and the categories of thought that it creates (Durkheim 1915, p. 20, n. 3; emphasis added).

While Durkheim's contemporaries were giving up arguments from origins and turning to pragmatist and propositional theories of knowledge, Durkheim (1915, p. 28; emphasis added) believed that something close to a classical notion of validity could be salvaged by securing a social origin for the categories: "If the *social origin of the categories* is admitted, a new attitude becomes possible, which we believe will enable us to escape both of the opposed difficulties [empiricism and apriorism]."

CRITICISM OF DURKHEIM'S ALLEGED ARGUMENT
FROM INTROSPECTION

Durkheim argues that participants in certain enacted practices have a direct experience of what he calls "moral force" and that this experience, which he often refers to as "internal," is the origin of the six categories of the understanding. Schmaus dismisses the whole argument on the grounds that "the experience of 'moral forces' fares no better than any other kind of *internal* experience as a justification of the categories" (emphasis added). Schmaus seems to think that the experience of moral force entails a justification from introspection of the sort ruled out by Hume. The classical argument from introspection turns on the belief that ideas that cannot have valid causes in external experience may have valid causes in the internal experience of bodily processes or consciousness. While Schmaus acknowledges that Durkheim accepted Hume's argument that the source of the idea of causation cannot be found in *external* experience of natural events, Schmaus believes that Durkheim overlooked Hume's argument that these ideas could not be found in our *internal* experience of the operation of the will. Schmaus writes, "Rawls appears to be unaware that Durkheim's attempt to ground causality in the experience of moral force leaves him open to David Hume's objection that one cannot perceive these forces or causes but only their effects."

There are four problems with Schmaus's claims in this regard: *First*, he interprets Hume as having made a special argument against introspection. However, what Hume (1975, p. 78) argues is not that there is a special problem with perceptions of internal states, but that the perception of *internal* states faces the same problems with regard to inferring relations between things, relations such as causality, as the perception of *external* states. In our internal states, causes still precede effects, as they do in external perception, and the cause and effect relation still cannot be perceived in a single moment. Therefore, even in introspection, an inference must be made by the mind to join cause and effect. Hume did not argue that there is something special about internal states that would make it impossible to perceive causality as Schmaus suggests but, rather, that there is nothing special about internal mental states in this regard and that internal states pose the same problem with regard to perceiving causal relations as the perception of external states of affairs.

Second, Schmaus interprets Durkheim's argument that the experience of causality is a direct internal impression, to mean that the idea of causality has been derived from a reflection on the internal workings of the mind. However, Durkheim did not make an argument from introspection. Durkheim argues that the primary reality giving rise to perceptions of

causal relations is external. When Durkheim says that the "feeling" of moral force is directly available internally as an impression of social forces enacted in rituals, he is not suggesting that through an introspective reflection on the relations between different internal mental states persons can validly infer the idea of moral force. He accepts Hume's argument in this regard (Durkheim 1915, pp. 406-10). What Durkheim argued is that moral force is available to immediate experience as a feeling created by the ritual and shared by all members of the group. He is employing a copy theory of ideas much like Hume's here. *The difference is that ideas are a copy of mutually created social relations not of naturally occurring individual phenomena.* Durkheim argued that the true purpose of ritual is to create these feelings and that the feelings are direct experiences of moral force. The relationship of necessary connection, he argues, can be directly perceived on certain social occasions and thereby comes to exist in the mind as an empirically valid impression of an external relation. There is no question of these impressions being the result of an operation of reflection, or introspection, on experience.

Third, Schmaus interprets Hume's argument regarding inference to mean that persons can perceive only effects but not causes. As a consequence, he argues, we cannot have internal impressions of causes, as Durkheim argues, but only impressions of effects. However, Hume does not argue that the objects or events that we refer to as "causes" cannot be perceived while "effects" can. When the causal inference is between physical objects, rather than invisible forces like "gravity" or mental states, both the object that we refer to as the "cause" and the object referred to as the "effect" can be perceived and would create in the mind impressions available to introspection. What we cannot perceive is the causal connection itself or that the thing being perceived is a "cause." Nor for that matter could introspection reveal that something is an "effect," which entails that it has a cause. The objects that we refer to as "causes" and "effects" are perceived as separate events, and their status as "causes" and as "effects" must in both cases be established by inference.

Fourth, Schmaus seems to feel that there is a problem with the idea that an external event (social practice) causes some internal event (the impression of moral force), and in parts of his text he treats the causal relation as the connection between the two, in which case, he suggests, the relationship itself needs to be inferred. However, when Hume says that persons have impressions of particulars available to introspection, he does not say that they are not real impressions, because the objects that caused those impressions are not also in our brains. If we are measuring Durkheim against Hume, emotion as an impression of an external social force should be accorded the same status that an impression of something external had for Hume.

SCHMAUS CLAIMS THAT DURKHEIM LACKED FAMILIARITY WITH HUME'S EMPIRICIST CRITIQUE

According to Schmaus, "Durkheim's knowledge of Hume was in fact rather incomplete. Rawls is just wrong when she says that 'Durkheim uses Hume's arguments with great sophistication.'" Schmaus's argument in this regard takes two forms. *First*, he argues that Durkheim displays an ignorance of Hume's argument that makes it clear that he did not understand it; and, *second*, he argues that Durkheim could not have had knowledge of Hume's arguments because, as a student of French philosophy, Durkheim had available to him only a very strange "spiritualist" version of Hume popularized in French philosophy textbooks.

With regard to the first point, Durkheim would have to understand three things. (1) He would need to understand Hume's argument that general ideas, including the idea of causality, could come only from inferences drawn from collections of particular sense impressions. For Durkheim, the problem with locating the origin of general ideas in perceptions of natural events is that perceptions are particular while concepts are general; perceptions are fleeting, unorganized, while concepts persist over time and impose an organization and meaning on things. He argues that it is not possible for a collection of particulars to be equivalent to any general idea. The general idea is not present in any of the particulars. It must be added to them by the mind and as such is merely a creation of the mind (Durkheim 1960, p. 34; 1915, pp. 26, 27, 31 n. 22, 171-73, 270, 407, 411, 480-81, 483, 490). This is a restatement of Hume's problem and recognizable as such.

(2) Durkheim would also need to understand the argument, articulated by both Hume and Kant, that the categories of space and time cannot be conceived without divisions. The classical empiricist dilemma with regard to space and time, that persons cannot form notions of quality or quantity without forming notions of degree, introduced by Hume in the *Treatise* (1975, p. 18), appears in Durkheim's (1915, pp. 22, 492) text in several places: "The mind cannot form any notion of quantity or quality without forming a precise notion of the degrees of each." For Hume, this argument presents an insurmountable circularity. For Durkheim, it does not, because he distinguishes the ability to perceive divisions in natural time and space from the perception of social or moral time and space. Only the latter give rise to valid categories.

(3) Further, Durkheim would have to understand Hume's critique of the argument from introspection. Durkheim discusses the problem with the argument from introspection with regard to causality at some length in the section on imitative rites. He writes, "In the first place, it is evident and recognized by all that it [causality] could not be furnished to us by

external experience. Our senses only enable us to perceive phenomena which coexist or which follow one another, but nothing perceived by them could give us the idea of this determining and compelling action which is characteristic of what we call power or force" (Durkheim 1915, pp. 406–7). General ideas furnished by introspection on the operation of the will would not work either. They would be entirely personal and therefore incommunicable. For Durkheim, the answer lies in socially created, shared, and communicable experiences that are internalized. According to Durkheim (1915, pp. 407–8), while the idea of force "can come only from our internal experiences; the only forces which we can directly learn about are necessarily moral forces." These are "impersonal forces" "coming from life together" that are directly available to internal experience. These impersonal forces fulfill Durkheim's (1915, p. 408) "two conditions," being experienced internally and as a collective external creation at one and the same time.

Second, in spite of the textual evidence to the contrary, Schmaus claims that because Hume's works were not readily available in France in Durkheim's day, Durkheim's familiarity with Hume was limited to several popular textbooks in philosophy representing the "official spiritualist philosophy" of the educated elite in France. As a consequence, Schmaus faults me for not considering the context of French philosophy in which Durkheim wrote. I understand the importance of considering theoretical works in their cultural and historical context. However, there are certain dangers in allowing a social context to determine the meaning of arguments. In Durkheim's case, the history of ideas method tends to identify Durkheim's argument with positions, like pragmatism, and French spiritualist philosophy, that he argued against. The method would tend to have this effect on the work of any thinker who differed significantly from their contemporaries.

There are problems with Schmaus's analysis of the French intellectual context in any case. Could French philosophy really have been ignorant of Hume? Most philosophers, whether they have agreed or disagreed with Hume, and whether they have thought he was a skeptic or not, have understood at least the basics of his argument, that because general ideas cannot have origins in the direct perception of natural events, they must be inferred. This argument had become something of a philosophical staple by Durkheim's day, and he makes frequent references to it in both *The Elementary Forms* and the lectures on pragmatism. In addition, pragmatism and Kantian philosophy, with which Durkheim was very familiar, were both in an important sense reactions to Hume's argument. It is hard to see how French philosophers could have embraced James's pragmatism, as they did, and challenged Hume's argument with regard to intro-

spection, as Schmaus claims, without understanding at least the basics of Hume's dilemma.

SOCIAL VERSUS NATURAL CAUSALITY

For Durkheim, the distinction between social and natural causes is key. Social causes are a product of group cooperation, to be clearly distinguished from the perception of relations between "things." For Durkheim (1915, p. 362) the categories, which could not be derived from relations between things, can be derived from the moral forces inherent in social relations "awakened in us by the spectacle of society, and not of sensations coming from the physical world." Schmaus completely dismisses Durkheim's distinction, claiming that "in his account of the notion of causal power, Durkheim appears to have been unaware that Hume's arguments would imply that we could not perceive social causes at all, not even obscurely."

However, Hume always analyzed causes from the standpoint of the individual. His argument has nothing to do with whether the sort of *social* causes Durkheim rests his case on can be perceived. Hume's argument is essentially irrelevant in this regard. Durkheim's distinction between social and natural causes opens up a new avenue of philosophical investigation, which he referred to as sociology. Durkheim argues that social forces, which are external, can, unlike natural forces, be directly perceived as emotions, which are internal. He argues that certain participatory social practices have a status very different from natural events and that a classical epistemological argument (of the origins form) that focused on social practices instead of on the individual perception of natural events, does not fall prey to the same objections. If the focus on the individual and the individual perception of natural forces is the cause of the problem, the focus on the perception of social forces, that can only be achieved in assembled groups, is Durkheim's solution.

That there is a difference between the social and the natural has been, from the beginning, one of the key differences between sociology and philosophy. It was a point that Marx, Weber, and Durkheim all insisted on. Marx argued that the inherent relatedness of all social phenomena was one of their distinguishing qualities. Therefore, for Marx, any method that treated mutually related and mutually defining social phenomena as separate and distinct was reifying. Marx's argument is not so different from the basic premise Durkheim is arguing for. The relatedness of social practice and moral force and of social practice and participants is inherent and real. It is not an inference added to or drawn from empirical reality after the fact by the human mind. Social forces are in an important sense

only relations. Although they are in some sense constituted by sounds and movements and in that sense empirically real, they must be constructed through social enactment. The sounds and movements only have significance and create effects when they are enacted as such by a group. Therefore, unlike natural forces, social forces can *only* be perceived as relations. With social phenomena, it is the individual that exists only by inference. The relation is primary and real.

In the lectures on pragmatism, Durkheim (1960, p. 67; emphasis added) reaffirms his emphasis on the inherent relatedness of social forces, arguing that "phenomena must not be represented in a closed series: things have a 'circular' character, and analysis can be prolonged to infinity. This is why I can accept neither the statement of the idealists, that *in the beginning there is thought*, nor that of the pragmatists, that *in the beginning there is action*." Durkheim, like Marx, insists on the dialectical or circular character of social relations. It is the primacy of *relations* in the social realm that makes the experience of relations a logical and a practical possibility. The great dichotomy between structural/functional and dialectical sociology does not begin with Durkheim.

CATEGORIES, REPRESENTATIONS, AND THE PROBLEM OF UNIVERSALITY

Because he thinks of the categories themselves as a priori, Schmaus argues that there is a difference between a given category and its collective representation. He asks how the category of causality could be inherent in the event? It is only possible, he thinks, that the event provide a collective representation of causality not that it creates the category. Schmaus says that I ignore this distinction even though, he claims, I am aware of it. Schmaus confuses two issues. For Durkheim, the experience of the categories in social practices comes first and creates the categories. After they have been created, collective representations may develop and attach themselves to the categories and the experiences that create them. It is not an issue of whether or not there is a distinction between the two but rather a matter of which comes first.

Schmaus begins with collective representations from which he claims Durkheim derived historically and socially contingent categories. Durkheim begins with the experience of categories in assembled groups and makes collective representations in general a derivative. The social process, not the idea, is the primary phenomena. According to Durkheim (1915, pp. 431–32), it is only in assembled groups that the moral forces corresponding to the categories can be felt. "The essential thing is that

men are assembled, that sentiments are felt in common and expressed in common acts."

Schmaus says that, on my view, the categories could only be mental contents: the result or copy of something external. "If the categories are supposed to be grounded in the experience of moral forces, it is difficult to see how she can distinguish the categories from sorts of representation, idea, or mental content." While for Durkheim, I believe they composed more of an emotional content than a mental content, so to speak, I do not in general see the problem with this. While from a philosophical perspective such an argument might appear to introduce hopeless contingency, it is not a problem from Durkheim's standpoint. Because the social feeling experienced has a common origin, the mental content is shared. The rite enacts the relation of causality, the direct experience of which is a "feeling" of causality. Once experienced, the categories are an embodied mental content, although they must be continually experienced in order to remain "vivid," which is why rites must be continually reenacted.

Schmaus claims that "to distinguish the categories from their collective representations," as he does, "is not necessarily to leave the categories themselves as some sort of free-floating entity." He believes the problem may be solved by correlating the categories with social functions. He says that "If one insists that these functionally defined concepts be represented somewhere, one could think of them as concepts developed by a community of social scientists and philosophers for the purpose of interpreting the collective representations of other cultures."

The problem with this interpretation is that it contradicts Durkheim's argument in every possible way. For Durkheim, the categories do not exist unless experienced. They cannot be given to a society by a group of social scientists and philosophers. Concepts given in this way would have no empirical validity. Because Schmaus believes that Durkheim remains a classical rationalist, he does not realize that Durkheim is arguing for the development of reason itself out of the experience of social practices. Categories of the understanding cannot be justified in terms of their social functions and at the same time have empirical validity.

Durkheim offers a very different sort of functional argument from the one Schmaus suggests. Schmaus says that Durkheim's argument is functional because "it appeals to the purpose of imposing obligations on the members of a society." But, Durkheim's account explains why societies can only exist where the rituals that create the categories have developed. *It is really the reverse of what Schmaus claims.* Durkheim argues that the purpose of religious rites is to create those emotions that are necessary in order for persons to experience the categories (Durkheim 1915, pp. 431–32, 469). The obligation presumes that performing the rite will create the

necessary emotions in the assembled group and has no value in its own right. Durkheim (1915, pp. 431–32) says that these emotions “depend upon the fact that the group is assembled, and not upon the special reasons for which it is assembled.”

Without the categories there is no reason, and without reason there is no intelligibility and no society. Therefore, society needs the categories and must insure that the social practices necessary to create them are performed. *It is an obligation imposed on societies by the need of social organization for reason.* Groups that do not force members to perform the necessary rites do not develop into societies, because their members do not develop categories of the understanding. If persons need to perform rites in order to develop reason, then society has to make sure the rites are performed or society will cease to exist. The categories are *not* created by this obligation, the obligation itself only exists because the categories can only be created by the *feelings* that result from *actually enacting* the practices in assembled groups.

CONCLUSION

Durkheim, contrary to what Schmaus supposes, attempted to *maximize*, not minimize, the contribution of social processes to the categories. He argued that discovering an origin for the categories in the group experience of social practices secured both their necessity and their validity. That Durkheim spent the last years of his life, in both *The Elementary Forms* and in his lectures on pragmatism (1960), offering a sociological alternative to empiricism, in its then popular form as pragmatism, is not unimportant. In the apparent absence of such an alternative, and in the wake of the demise of classical rationalism, various positions that base the limits of knowledge on arbitrary conventions of language or social action, such as those assumed by Schmaus, have come to dominate contemporary social theory and philosophy. Durkheim made an argument with the potential to alter the epistemological landscape in both his own day and in ours. Schmaus has not recognized the novelty of Durkheim's position. Instead, he reintroduces the arguments generally leveled against Durkheim, continuing the refrain that Durkheim was not aware of these objections. Not only was Durkheim not ignorant in this regard, he was well aware that foundationalist arguments had never before succeeded, and he argued that they failed precisely because they did not recognize the essential role played by social processes in creating the human faculty of reason. Rather than seeing the prior failure of such arguments as evidence that Durkheim would not have tried to make a foundational argument, as Schmaus does, we should see that Durkheim saw himself as a

pioneer in this regard. That is what makes his challenge to the philosophical debate over epistemology so interesting and important.

REFERENCES

- Durkheim, Émile (1912) 1915. *The Elementary Forms of the Religious Life*. Chicago: Free Press.
- . (1913–1914) 1960. *Pragmatism and Sociology*. Ann Arbor: University of Michigan Press
- Hume, David (1739) 1975. *A Treatise of Human Nature*. New York: Oxford University Press.

ANNE WARFIELD RAWLS

Wayne State University

Review Essay: Chasing Alfred Chandler

Big Business and the Wealth of Nations. Edited by Alfred Chandler, Franco Amatori, and Takashi Hikino. Cambridge: Cambridge University Press, 1997. Pp. xii+596. \$59.95.

Neil Fligstein

University of California, Berkeley

It is undeniable that economic growth in Western Europe and the United States (and, since 1960, in parts of Asia) has outstripped the rest of the world. The causes of this uneven growth are the core subject matter of economics, political economy, and now, thanks to Alfred Chandler and his colleagues, business history. Why should we care about this? For the past 20 years or so, the field of economic sociology has been flourishing. We have excellent theoretical and empirical studies of the myriad ways in which social relationships and governments have shaped the development of particular organizations and markets. We even have the beginnings of a comparative study of capitalist societies and societies in transition from socialism to market economies.

But economic sociologists have failed to demonstrate systematically why various social arrangements are or are not consequential for economic growth and development. We have skirted the question of efficiency and more or less fallen back on the argument that the world is more contingent than economics would allow. Whether this means that there is more than one way to organize markets or, alternatively, that market processes do not strongly select organizational forms is a question that remains elusively beyond our grasp. Chandler has upped the ante for sociologists by claiming to explain the differences in economic growth among advanced capitalism societies and by arguing for the importance of the corporation in that process. He is willing to say that there are different paths to economic growth, but that some are better than others. I will argue that Chandler and his colleagues do not, in the bulk of their case studies, provide clear causal evidence for Chandler's central assertion. Indeed, their evidence suggests that sociological approaches that compare societies and add more analytic constructs as causal factors are much more effective at uncovering the truth about economic growth. But in spite of this, Chandler et al.'s relatively clear statement of the questions and analytic concerns that confront all research on the social organization of economies offers a positive model by which to evaluate our own work.

The theoretical essays in the book make an important claim: the best way to understand the varying performances of the world's economies

© 1998 by The University of Chicago. All rights reserved.
0002-9602/99/10403-0010\$02.50

in the past 150 years is to study the large industrial enterprise. Chandler asserts that wherever the large firm emerged to organize productive assets (i.e., raw materials, machinery, human resources, and technical knowledge), economic growth followed. Where firms were slower to develop, or where firms failed to develop altogether, economic growth did not materialize. This has happened in two phases. First, Chandler and Hikino argue that during the end of the 19th century, large industrial firms emerged to take advantage of economies of scale and scope presented by the possibilities of new technologies. The most important problems that managers and owners needed to solve was how to control inputs and outputs to gain the advantages of the new technologies (i.e., to exploit economies of scale). Second, they also found new ways to use the products of such processes to gain economies of scope. They began to realize that by doing this they created the possibility of building whole new markets. After World War II, the large corporation applied this lesson even more systematically by investing in research and development and human resources to build long-term capabilities. Chandler's assertion is that successful firms created their own new products and developed technologies themselves. Where these successful firms existed, economic growth has been high; where it has not existed, economic growth has been impeded. Managers and corporations are the heroes in Chandler's stories because they are the professionals who, by continuously learning and experimenting, create the wealth of nations.

The central causal problem of Chandler's approach is isolating the effect of the existence of the capitalist firm on economic growth. An economist would say that markets produce wealth by getting actors to allocate scarce resources to maximize the value of those resources through the price system. Economic growth results when free markets are able to effect investment decisions about land, labor, and capital in the most productive way. Firms are endogenous to the market and arise as responses to opportunities to create wealth. A Marxist would say that the search for profits causes wealth and that technology increases labor productivity, which in turn produces more surplus value. Economic growth, then, is spurred by the dynamics of capitalist accumulation. Many versions of "new" institutionalist arguments would focus on creating a system of institutions that could organize property rights, legally control competition, and make rules to favor exchange. Finally, political economists of all varieties would emphasize the role of governments in providing the right kind of services for economic growth, such as a predictable political process, an adequate infrastructure, stable currencies, and an educated labor force. For these other theories, large firms *and* economic growth are outcomes of the political and economic organization of societies.

To demonstrate that large corporations are the real engine of economic growth, one would have to conduct econometric studies that contain measures of the conventional sort suggested by other theories. Convincing research would also have to look at the ways in which corporations really differ across societies. This is not what any of Chandler's colleagues have

done. Instead, their strategy is to look at more and less developed societies (say, the United States and Spain) and show whether or not the "good" kind of management hierarchies emerged. The problem is that most of the case studies in this volume show the great diversity of firms: they vary by size, by labor relations regimes, by the number of products they produce, by the degree of vertical and horizontal integration, by ownership, and by which functions tend to dominate top leadership. Moreover, most of the case studies point to the pivotal roles played by governments in dealing with problems of social, economic, and institutional instability, and in particular, in helping resolve problems of worker-manager conflict. War, depression, and imperialism play a part in the relative success of governments that work to produce stable conditions for economic growth, and in most of the papers, the emergence of large firms seems endogenous to such processes. Even Chandler admits that the postwar investment in technology in the United States was an outgrowth of the war economy (p. 83), not an endogenous corporate search to create new markets. Of the 11 chapter studies of individual societies included here, most provide evidence that does not support Chandler's view that economic growth is most fully realized when managers focus on vertical integration, economies of scale and scope, and the production of new technologies. Put simply, Chandler's heroic managers do not find stages to work their magic unless they are properly situated in stable, non-rent-seeking political economies. Even then, they organize in a great many ways that reduce to finding new product markets by making investments in technology in only the most superficial sense.

The last five chapters of the book take up these questions more systematically and raise important objections to Chandler's thesis, some of which I have discussed above. Giovanni Dosi makes many criticisms, among them—and perhaps most important—that it is difficult to partition the variance in economic growth to various factors. Takashi Hikino focuses his discussion on forms of ownership, how they differ across society and the theoretical and empirical results obtained on their relative effects on economic growth. He suggests that Chandler's analysis ignores these differences. William Lazonick and Mary O'Sullivan focus on how skill formation in societies has a lot to do with their educational systems and with the role of workers in firms. They disagree with Chandler that the only important skills are those that emanate from managers. They suggest that American managers generally see workers as "costs to be reduced" (p. 520) and that this seems to be antithetical with skill enhancement or creating new products. Thomas McCraw examines some comparative evidence on the role of governments in economic development and concludes that there is evidence that governments can matter for economic growth. Finally, Jeffrey Fear reviews what he calls the "cultural conception of the firm" and concludes that Chandler's argument makes it hard to take into account the very different trajectories firms pursue.

Alfred Chandler the organizational theorist and Alfred Chandler the historian often come into conflict with one another. The opening chapters

of all of Chandler's books offer clear and simple hypotheses about what makes organizations successful. But, a close reading of the historical materials in his books provides ample evidence that other factors are easily as pivotal to the outcome as the one he favors. This book is no different on that front.

But, even if Chandler cannot prove his thesis, he has laid down the lines of debate that are important if we wish to make sense of where firms and capitalism are going. Economic sociology, as a field, has failed to grapple head on with many of these themes. Capitalism has produced economic growth beyond anyone's understanding. While development remains uneven, it has occurred in many places. The organization of capitalist markets and firms across time and space presents many puzzles. On the one hand, there appear to be distinct national styles of organizing firms, and many of these continue to appear to be viable. On the other, more markets worldwide are becoming more integrated. We have not done a good job of thinking about why there has not been more convergence in organizational forms. A stable political and economic environment and a system of employment relations are as important to the existence of corporations as the profit-generating activities they undertake. Alfred Chandler may be wrong that the only engine of economic growth is the managerially led large corporation. But an economic sociology that fails to offer a viable alternative account is one that will end up chasing Alfred Chandler.

Book Reviews

Poverty, Ethnicity, and Violent Crime. By James F. Short, Jr. Boulder, Colo.: Westview Press, 1997. Pp. xii+244. \$49.00 (cloth); \$16.00 (paper).

Robert J. Bursik, Jr.
University of Missouri, St. Louis

Few people have had as much influence on the substantive and methodological character of contemporary criminology as James F. Short. But his centrality to the discipline is not solely a function of the major contributions that he has made during his prolific career. Rather, he personifies the link between the classic Chicago schools of criminology and the modern frameworks whose intellectual roots are firmly embedded in those traditions. Despite the important theoretical differences between the various schools, the shared hallmark of all the Chicago approaches was a deep interdisciplinary appreciation for the multiple contexts through which behavior develops. *Poverty, Ethnicity, and Violent Crime* is an exemplar of this orientation.

The central theme that unifies the material in this book originally was presented by Short and Albert Cohen in a chapter on juvenile delinquency in the first edition of *Contemporary Social Problems* (edited by Robert K. Merton and Robert A. Nisbet [Harcourt Brace Jovanovich, 1961]) and continually has been refined by Short in a number of subsequent publications. Succinctly stated, he argues that a full understanding of the associations among poverty, ethnicity, and violent crime only is possible through a synthetic understanding of processes that operate at three levels: the individual (biological and psychosocial processes), the macrosocial (characteristics of organizations, social systems, social structures, and cultures), and the microsocial (the situationally specific unfolding of an interactional event between two or more parties).

The book pursues the development of this theme in four basic sections. The first (chaps. 1–3) not only introduces Short's synthetic framework but also provides an encapsulation of historical research on violent behavior and an introduction to its measurement. While most of this material will be familiar to criminologists, it is essential reading for those outside the field who desire more than a superficial understanding of violent crime. At the least, the historical data will shatter any remaining illusions that some golden age of tranquillity existed in the past. In addition, there are excellent summaries of the trends and distributions of violent crime that have characterized various socioeconomic and ethnic groups.

The substantive core of the book begins in the second section (chaps. 4–6) where, drawing from the findings of a very comprehensive body of

Permission to reprint a book review printed in this section may be retained only from the author

the research conducted by the author and others and focusing in particular on the activities of youth groups and gangs, Short addresses the degree to which the effects of poverty and ethnic status on violent behavior are mediated by the local community contexts in which such behavior occurs. This section of the book easily could be considered to represent volume 2 of his classic *Groups Processes and Gang Delinquency* (with Fred L. Strodbeck [University of Chicago Press, 1965]) and represents Short at his best. In particular, his illustrations of how the convergence of processes operating at all three levels leads to complex nonrecursive relationships among social capital, human capital, and crime are insightful and provocative.

The theoretical explanations for these empirical patterns are more fully considered in the third section (chaps. 7–9). I admit to being somewhat let down by this part of the book. There is no question that his discussion of the theoretical perspectives that have been developed at each of the three levels draws from an astonishing range of material. However, there is little effort to integrate these models into a logically consistent package, as he attempted with the empirical literature he reviewed in the previous chapters (Short himself notes this problem on p. 177). This failure should not be judged especially harshly, for throughout the book, Short notes that such endeavors are constrained significantly by our lack of understanding of many aspects of violent behavior and the tentative nature of what we think we do know. Nevertheless, I think that Short shows convincingly in the second section that a social and human capital orientation provides a very promising basis for an initial integrative endeavor. This potential is reinforced in the third section in a passage entitled “Extending the Theoretical Implications of Capital” (pp. 172–75). I look forward to his future work, which will more intensively develop these themes.

The book closes with a one-chapter section on the policy implications of his three-tiered model. This material basically represents a call to action on the part of contemporary criminologists. In fact, in a quotation from William J. Wilson that is used as the final sentence of the book (p. 204), it is noted that if criminologists do not get involved in policy debates, “decisions will be made and policies will be formulated anyway—without their input.” The degree to which Short links modern criminology to the ideals of the old Chicago schools may not be more clearly illustrated than in this strong commitment to the social responsibilities of the academic community. I can think of few legacies that could be more valuable to pass on to future generations of social scientists.

Violence and Childhood in the Inner City. Edited by Joan McCord. New York: Cambridge University Press, 1997. Pp. xiii+334. \$59.95.

Darnell F. Hawkins
University of Illinois, Chicago

The boldness and self-confidence (or self-deception) that allowed earlier generations of social scientists to tackle complex social phenomena and proclaim that their own theories and modes of empirical analysis offered *the* answer to significant research questions are now largely a thing of the past. As sociologists, we have come to see the value of incorporating into our research designs not only those diverse perspectives found within our own discipline but also those associated with all disciplines that study humans and the social world. In almost every area of social inquiry, it has become fashionable today to posit notions of multiple causes, mediating and moderating factors, etiological contingencies, and proximal and distal effects, all of which allude to the benefits to be gained from multidisciplinary, multimethod approaches to the study of social problems.

This excellent volume on childhood and violence in the nation's inner cities by Joan McCord and her colleagues rides the waves of the current penchant for theoretical and empirical analyses that incorporate a diversity of research foci and points of view. Though broad in its coverage, the collection of essays offers a cohesive and well-conceived portrait of current thinking regarding the nature and consequences of inner-city violence. The volume grew out of a series of meetings among the contributors that were supported by the Harry Frank Guggenheim Foundation, and it contains seven chapters.

The themes explored by these chapters include (1) "street culture" and other manifestations of subcultural differences between inner-city and "mainstream" youth (Elijah Anderson, also Robert Sampson and Ronald Slaby), (2) community and neighborhood disorganization and its effects on childhood and adolescent development (Robert Sampson, also Terrie Moffitt), (3) the effects of American cultural values, urban life, and racial discrimination on the inner city (Joan McCord, also Robert Sampson and Elijah Anderson), (4) the neuropsychological determinants of aggressive and violent behavior (Terrie Moffitt, also Ronald Slaby and Nancy Guerra), (5) victim and bystander effects and other psychological mediators of violence among urban youth (Ronald Slaby, also Nancy Guerra), (6) child abuse and its prevention among inner-city populations (Tony Earls and Jacqueline Barnes), and (7) factors affecting the implementation of inner-city violence reduction programs (Nancy Guerra, also Tony Earls and Jacqueline Barnes).

Both as individual chapters and as a collection, this book is very impressive. The book contributes to and builds a strong case for the wisdom of current attempts among violence researchers to combine individual, situational, and macrostructural levels of analysis and theory. Further, each of the chapters fully succeeds in convincing the reader of the merits

and importance of the approach taken by the author(s) for understanding the problem of violence and childhood in the inner city. By doing so, each lends support to McCord's assessment in the book's preface and introduction of the value of multiple opinions regarding the causes of inner-city violence and the changes necessary to reduce it. McCord's chapter also places the contemporary problem of inner-city violence in its proper historical and cross-national context.

As a researcher who has examined racial and class differences in rates of violent offending and victimization, I was particularly impressed by Terrie Moffitt's thoughtful discussion of what we currently know and, importantly, do not know about the neuropsychology of violence, as well as the link between neuropsychological deficits and the inner-city neighborhood context. Her well-written review of a very complex and often jargon-laden literature makes clear to the average nonspecialist reader just what research questions drive this area of inquiry and what conclusions can be reached on the basis of extant knowledge. She demystifies an area of research that is characterized by exaggerated claims by some researchers of the strength of the association between neuropsychological deficits and violence and the claims of others that such research has nothing to contribute to the study of inner-city violence.

My one area of criticism of the book stems largely from the feature that is its greatest strength—its incorporation of multiple research approaches and suggestions for the remedy and amelioration of the inner-city violence problem. On one hand, this collection is an anthem to our enlightened attitudes toward the need for a diversity of viewpoints in the study of social problems. As one who supports such diversity, that feature contributed to the book's appeal for me. On the other hand, upon completion of my reading of this informative and stimulating collection, the "economic determinist" in me longed for a simpler and more straightforward analysis of inner-city childhood violence—one that would offer more finality and definitive conclusions. Where is Karl Marx when we need him? Where is the Émile Durkheim who acknowledged the effects of biological, psychological, and contextual forces as explanations for variations in deviance and violence but came down squarely in favor of social structure as *the* determining factor? Maybe I should reread Robert Merton.

Violence and Truth: A Rajasthani Kingdom Confronts Colonial Authority. By Denis Vidal. Delhi: Oxford University Press, 1997. Pp. ix+244. \$29.95.

Frank de Zwart
Leiden University

Violence and Truth is an interesting study on the impact of colonial rule on society and politics in Sirohi, a small Hindu Kingdom in West India. The book focuses on the late 19th and early 20th centuries, though it

often considers traditional society. Vidal writes for an audience of historians and sociologists. His central thesis, echoing structural functionalism, is that caste-specific "patterns of protest"—a concept closely resembling Charles Tilly's "repertoires of collective action" (*From Mobilization to Revolution* [Addison-Wesley, 1978])—traditionally functioned to maintain the balance of power in Sirohi. The British colonial government suppressed these protests because they challenged its monopoly of violence and thus deeply affected the social structure.

Vidal presents a vivid picture of caste-specific protest repertory. Violent revolt was a standard means of protest for Rajput nobles (holders of revenue-yielding estates) in conflict with their king. Conflicts between the king and nobles primarily concerned lineages within the dominant Rajput clan, but other groups in society were also involved: villagers whose life and property were targets of attack, tribals who lent military assistance, and Brahmins and Charans (bards) who mediated and used their own special repertory item, threats of collective suicide. Indirectly, conflicts also often affected Jains (trading and merchant castes) whose specific way of protest was to migrate. Vidal stresses that these traditional patterns of protest and the values they represented were vital for regulating the way society functioned (pp. 5, 55, 154, 217). In confrontations between the castes, relationships were redefined, and "principles on which the balance of society rested" (p. 217) were put to the test. Without confrontations, the caste system could not function.

The system of values upon which caste society in Sirohi rested, Vidal argues, was fundamentally different from that which informed British legislation. In precolonial Sirohi, violence was "invariably assessed according to who committed [it] against whom" (p. 15). British law criminalized all violence, however, and offenders were tried in court. Consequently, nobles were forced into other, less effective forms of protest—such as court cases—and lost power to the king.

British values also affected the status of bards. Bards (belonging to the Charan or Bhat castes) in traditional Sirohi were specialists in genealogy, biography, and accounts of historical events. They were an important caste, because through their mediation, the "collective identity of the ruling [Rajput] clans was defined and perpetuated" (p. 93). The status of bards derived from their acknowledged competence to express truth. Assessment of this competence, however, depended not so much on the "supposed equivalence of their narratives to actual facts, as on the sacred nature of their inspiration" (p. 102). The British legal system, however, became increasingly relevant in settling conflicts in Sirohi, and it showed total incomprehension of the different criteria for truth in the native cultural tradition. Accepted truth in British institutions was based on empirical facts. As a consequence, the bards lost their monopoly on truth, and their privileged status was undermined.

The British, Vidal concludes, promoted "homogenization of the standards underlying social usage" (p. 219). Acts of violence were no longer legitimate when committed by the "right" person but became simply

criminal; truth no longer depended on who told it but became objective fact. To put it short, British influence destroyed the moral and scientific relativism on which the traditional order in Sirohi rested.

Though this is an interesting thesis, Vidal does not elaborate on its relevance for social science theory. He writes that studying history is the best way to find out "where we have gone wrong and what we have overlooked so far in the sociology of India" (p. 8). Indeed, many issues of interest to social theory are touched upon, but none of them is pursued. Vidal's main thesis, for instance, echoes Max Gluckman's structural functionalist analysis of ritual and social order. There is no reference to this theme, however. Instead, Vidal names "modern" issues such as social change and essentialism but leaves unclear exactly which sociological mistakes he wants to correct. Flirtations with postmodern sociology (pp. 4-6), on the one hand, suggest that his enemy is the "essentialist" view of India's caste system, contributed commonly to the influence of Louis Dumont (*Homo Hierarchicus* [University of Chicago Press, 1970]). But Vidal's account of traditional caste society, on the other hand, with its caste-based occupation and status hierarchy of priests, warriors, and traders, confirms rather than defeats Dumont's view.

Vidal also raises the question as to why Gandhi's nonviolent protest pattern was so successful. From the analysis of Sirohi politics, he suggests, it follows that we should explain the success of nonviolence as an unintended result of British law and not, as some authors do, as a traditional Indian protest pattern. British law imposed a nonviolent protest pattern upon society; Gandhi made that backfire. The reader is led to think that this explanation represents Vidal's broader theoretical intentions, but next he is told that migration (another protest pattern) is linked to the scriptural Hindu value of "withdrawal," a "fundamental cultural factor in India, pervading its entire social logic" (pp. 217-18).

To some who are familiar with current debates on caste in India, finding these two explanations in one book is refreshingly eclectic. However, one wishes the author would have taken more trouble to make a point of that. The many interesting data he presents deserve clearer theory.

Racism in a Racial Democracy: The Maintenance of White Supremacy in Brazil. By France Winddance Twine. New Brunswick, N.J.: Rutgers University Press, 1998. Pp. xiii+175. \$49.00 (cloth); \$17.00 (paper).

Evandro Camara
Emporia State University

France Winddance Twine's recent ethnographic study of Brazilian race relations is a welcome addition to the literature. It is a clearly written, forceful, and thought-provoking work that should stimulate the ongoing, often volatile, debate on comparative race and ethnicity. I was especially intrigued by the author's attention to ideological and discursive patterns,

which are presented as sustaining mechanisms of racism in Brazilian social life. Unfortunately, for all of its merits, this work displays the familiar sins of the revisionist scholarship on comparative race and ethnicity. This significantly compromises its value.

We may begin with the title, which, in the spirit of revelation, calls attention to the existence of white supremacy in Brazil. This mainly restates the *generality* of a problem whose persistence worldwide as the legacy of centuries of Western European colonialism has long been recognized. It also expresses the analytical homogenization of racism, which assumes the universal validity of race as the cornerstone of social organization, and of a conceptual and analytical lingua franca, applicable to all kinds of interethnic situations. As a former European colony, Brazil, notwithstanding its romanticized popular and scholarly depictions as a racial paradise, expectedly displays this general aspect of ethnic inequality. But, the analysis should go on to identify and dissect the *particularity* of the problem so useful cross-national comparisons can be made.

The homogenizing tendency has rendered this work ethnocentric, at best, and at worst, a regular case of cultural colonization. The premises and language of biracialism operate here as a kind of Archimedean point wherefrom various aspects of race and ethnicity in Brazil are evaluated. The population is treated dualistically as consisting of white and non-white segments, a practice that negates the pattern of ethnic and racial structuration in Brazil—clearly a case of forcing square pegs into round holes.

Twine relies on the UNESCO data (a large body of ethnographic research done on race and ethnicity in Brazil in the early 1950s) to bolster her claim that racial prejudice creates significant social distance between “whites” and “blacks.” But she treats this material too briefly to benefit from its incisive examination of the complexities of race and class in that country. For instance, she appeals to Marvin Harris’s account of the difficulties of social assimilation of a “Negro councilman” in a small Northeastern town, but she fails to consider his observation that this individual was not handicapped solely because of his skin color but mainly because he was not “rich enough or educated enough to overcome being a Negro” (Harris 1952, p. 69; my emphasis). This situation is paradigmatic and underscores the fact that the “whitening of race” process, despite its associations with the overvaluation of aesthetic ideals and traits pertaining to European ancestry and with material privilege (as Twine correctly notes; p. 109), cannot be separated in Brazil from the assumption of transformability of racial status and universal social inclusion. Hence, the fluidity of racial condition (vs. its quasi-metaphysical fixedness in biracial social systems), which has inhibited the crystallization of ethnic minorities. The point is that, although the legacy of sociostructural disenfranchisement left by slavery still haunts “Afro-Brazilians,” it is countered by the possibility of changing racial status, of transforming group *difference* into *nondifference*.

Racial essentialism underlies the author’s discomfort with the national

ethos of *embranquecimento* (whitening) of the population and concomitant practice of *mestiçagem* (intermixture). She states that in Brazil the "one-drop rule" works in reverse (p. 113) in that any amount of "whiteness" leads to the absorption of the "black" into the "white" population, eventually "erasing" the former. This is supposedly detrimental because it prevents the "people of color" from having a separate existence and identity. One is led to wonder how the social inclusion of a group can possibly imply its marginalization. Moreover, is there not an assumption of ontological domination here, that is, of an essential *negritude* that is being denied, through the whitening process, to all but members of the "white elites"? If allowed to flourish, this condition would permanently sustain a "black" community, as in the castelike separatism of biracial systems. I scarcely think this to be a desirable outcome for ethnic minorities anywhere.

I do not wish to deconstruct the idea of race entirely. Where operative as the key regulative principle of social relations, race fulfills important bureaucratic and psychosocial functions. Still, I want to relativize it cross-nationally so as to neutralize its worst reifications.

The imposition of a biracialist *armamentarium* in this book has ridden roughshod over historical and cultural particularity and the fact that race and ethnicity must be addressed in the context of the social system. Aspects related to the dominant-minority cultural relationship, which are crucial for differentiating between interethnic models, were basically neglected. The Brazil that emerges from this analysis therefore is simply *another United States*—only an inferior, less progressive version.

Nationalizing Blackness: Afro-Cubanismo and Artistic Revolution in Havana, 1920–1940. By Robin D. Moore. Pitt Latin American Series. Pittsburgh: University of Pittsburgh Press, 1997. Pp. xii+320. \$45.00 (cloth); \$19.95 (paper).

Lauren Derby
University of Chicago

This is a meticulously researched and finely textured analysis of the blackness vogue that swept Cuba in the 1920s and 1930s and how it redefined the Cuban nation as inclusive of its mixed-race heritage. Written by an ethnomusicologist, the text focuses on music but pays close attention to popular theater and the fine arts as well. Indeed, such a multidirectional methodology is necessary since the "Afro-Cubanophile frenzy" Moore describes swept salon and dance music, poetry, painting, folklore, and ethnography alike. The book consists of a theoretical introduction and conclusion and six thematic chapters on popular theater, bourgeois and street carnival traditions, the origins of salsa music, the rumba craze, cultural nationalism, the folklorization of black culture, and the avant-

garde modernist literary canon. It presents the first thorough cultural history in English of a period of extraordinary creativity in the fine arts in Cuba, yet one that also gained momentum from Josephine Baker's primitivist Parisian revues, the popularity of jazz, and the Harlem Renaissance. By excavating many of the lesser-known artists, musicians, and writers of the period, Moore demonstrates that attention to Cuban literati greats such as Nicolás Guillén and Alejo Carpentier has caused scholars to overlook the variety of views among their rank and file followers. Moore is a master craftsman when it comes to documenting the multiple strands of debate within each movement, and he inspires with his attention to nuance and specificity of detail.

The organizing question of the study is how and why, given the prevalence of racism, Afro-Cuban themes came to dominate Cuban music, theater, and painting during this period, eventually coming to symbolize Cuban national identity itself. The author traces how cultural practices associated with blackness changed in meaning as they passed from the hands of rural musicians playing *son* guitar music in the 1890s, to the global audience of the rumba vogue of the 1920s, to the elite avant-garde modernist movement that paralleled the 1933 Cuban Revolution. Moore breaks new ground in his exhaustive stylistic documentation of previously unexplored themes such as Cuban minstrelsy and blackface, as well as the street music, *comparsa*, characteristic of popular carnival procession in turn-of-the-century Cuba. In charting the changing contours of Afro-Cuban artistic genres over time, Moore rejects the "African survivals" approach pioneered by Melville Herskovits in favor of a dynamic model in which the combined effect of new audiences, folklorization, and state repression dramatically altered the meanings of Afro-Cuban cultural forms and practices over time.

Moore describes how the 1920s global vogue for things Afro-Cuban perplexed and dismayed certain Cuban elites who wanted their nation to be represented overseas as "civilized" rather than "primitive." Just as Brazilian actress Carmen Miranda, much to her chagrin, was rejected by Rio high society upon her return from Hollywood stardom due to her type-casting as a sensual mulatta, the Cuban upper classes were mystified by the thirst for Afro-Cuban music in the United States and Europe since it challenged their positivist view of black culture as atavistic and of culture as a sign of civilization and progress. While recent Cuban scholarship has focused on tensions surrounding Afro-Cuban political participation in the postemancipation period, this is the first study to chart how racial ambivalence resulting from a contradictory desire to include Afro-Cubans within the nation while excluding them from effective political participation was expressed in the sphere of cultural production. However, Moore goes far beyond considering just how political struggles are simply reflected in the sphere of culture. He demonstrates how whites developed a mimetic desire to impersonate Afro-Cuban personae in song and theater—to actually become the black cornerstore peanut vendor in a minstrel show or the Santería priest—at the very moment when on the streets

of Havana police were raiding Santería shrines and confiscating ritual objects.

Given Moore's rich documentation of how racial meanings have been transformed through processes of cultural appropriation, the actual sociology of race relations, including those who fall outside of a racially dichotomous black/white framework, could have been highlighted more in his account. A bit more attention to the ways in which race and class intersected during this period, for example, might help explain some otherwise curious phenomena, such as the fact that middle-class Afro Cubans were some of the most vocal critics of Afro Cuban culture in the early 20th century, at times stridently calling for legislation banning black musical genres in public. One possible explanation might be that these Afro Cubans were seeking to claim an elite status position by censoring phenomena associated with poverty in a context in which idioms of race and class were deeply intertwined. Moreover, how did the contours of who was defined (and defined themselves) as black change from emancipation onward? For example, whites during this period gradually came to penetrate Afro Cuban cultural domains such as Santería houses, becoming practitioners, high priests, and even establishing their own *cabildos* or ritual houses, a process that created a gap between the ascribed race of certain social spaces and the participants themselves (Peter Wade, *Blackness and Race Mixture: The Dynamics of Racial Identity in Colombia* [Johns Hopkins University Press, 1993, p. 272]). Moreover, the perspective of the "mixed-race" population is given short shrift in this account.

In addition, one wonders what roles the United States might have had in shaping racial meanings in Cuba during this period of both repeated intervention and direct rule. The U.S. government in Haiti, for example, paved the way for the eventual *noirist* backlash led by François Duvalier by establishing Jim Crow segregation and buttressing the mulatto elite's political position. At the very least, the early Americanization of the Cuban elite—who by the late 19th century were staging their debutante parties in Miami—may have encouraged Cubans to see their nation through the prism of a U.S. racial lens, even if a substantial population of mulattos made it a quite different reality. Of course, one difficulty in treating Cuba during this tumultuous period is that the story of race is difficult to isolate from other contextual factors. For example, when the student union took over government with support from communists and rural labor in 1933, public life was regulated on many fronts as a fragile regime sought to reestablish civil order. Thus, this context of active civil strife causes one to question whether indeed municipal efforts to regulate *comparsas* or black carnival street parades were primarily or only secondarily about repressing a sign of blackness.

Moore does an outstanding job at finding a vocabulary for talking about race in a context in which it has a curious way of slipping in and out of focus. In stressing the extent of racial prejudice in Cuba during this period, however, at times he may be displaying a normative bias

toward a civil rights model of racial ideology that does not always fit the context. This is most evident in the conclusion in which he stresses the common ground between racial ideology in the United States and Cuba. Unlike the United States, Cuban national identity was implicitly perceived as ethnically homogenous, which explains why, as in Ecuador or Colombia or the Dominican Republic, to define oneself as black is to effectively abdicate membership in the nation (Ronald Stutzman, "*El Mestizaje: An All-Inclusive Ideology of Exclusion*," pp. 45–93 in *Cultural Transformation and Ethnicity in Modern Ecuador*, Norman E. Whitten, Jr., ed. [University of Illinois Press, 1981]). The fact that even Afrocuban artists performed in blackface in Cuba in the 1920s indicates just how much more complex are the multiple and contradictory meanings of race in Cuba than the models of "plural societies" assumed (Michael G. Smith, *The Plural Society in the West Indies* [University of California Press, 1966]). This is a fascinating study that makes a major contribution to the cultural history of race and national identity in the Caribbean, while at the same time establishing the highest of standards for Latin American cultural history, at the level of both rigorous documentation and theoretical breadth.

Inside the Revolution: Everyday Life in Socialist Cuba. By Mona Rosendahl. Ithaca, N.Y.: Cornell University Press, 1997. Pp. x+197. \$37.50 (cloth); \$14.95 (paper).

Sheryl L. Lutjens
Northern Arizona University

Inside the Revolution: Everyday Life in Socialist Cuba makes a welcome contribution to the literature on Cuba. Empirical research has long been difficult for foreign scholars, and Swedish anthropologist Mona Rosendahl's 18 months of fieldwork (1988–90) in a municipality she calls Palmera provide an unusually good basis for analyzing Cuban socialism. Keen on studying the everyday meaning of a "strong, unified, and hegemonic ideology" (p. 3), Rosendahl began with questions about the delivery and reception of ideology, including the interaction of socialist ideology, existing beliefs, and sociodemographic characteristics. The ethnographic analysis of her experiences living in the capital of Palmera informs an intelligent, sometimes comparative, and ultimately provocative local reading of crucial political issues of Cuban socialism. The epilogue based on short visits (1993 and 1995) during the crisis conditions of Cuba's Special Period is vital to Rosendahl's assessment of what the Revolution—and its ideology—has meant to ordinary people.

Rosendahl frames her study with "three approaches" that, together, are to uncover the dynamics of ideology (defined as a "set of ideals that deals with social and social relations," with explanatory, normative, and practical content; p. 3): official ideology based in Marxism-Leninism, ideology

in daily life as it is re-defined by collective memories and individual experiences, and the interface of official ideology with other "complexes of ideas" (pp. 3–4). It is with two such complexes—reciprocity and gender relations—that Rosendahl introduces her readers to the realities of socialist ideology and institutions in Palmera. Reciprocity is not new in Cuba, she notes, though the meaning of loans, barter, and gift giving is explained with reference to the socialist planned economy and the informal realities of consumption. Similarly contextualized with reference to history and a foundational *machismo*, gender relations are a second complex of ideas in which practices and shared beliefs in Palmera—ranging from marriage and motherhood to eroticism—fall short of official ideology's call for equality between men and women.

The analytical description of daily life is the pivot for subsequent examination of the Cuban Communist Party, official discourse, and the "revolutionary act." The Leninist tension between centralization and participation is key. Rosendahl describes "total centralism" ensured by an ideology-creating party with "ultimate power concerning all societal activities" (p. 81); revolutionary rhetoric is situational and criticism difficult; and being a good revolutionary means active participation and sacrifice. Observations of meetings of mass organizations and local government (Cuba's representative institutions), public speeches and rituals, and ordinary people's participation including voluntary labor, lead Rosendahl to some general conclusions. All Cubans are affected by hegemonic socialist ideology, though acceptance of it—and support for the Revolution—is conditioned by personal experience, position in the social hierarchy, and race; party membership signals the most noticeable differences in attitudes, and age, education, and gender the least. Rosendahl further concludes that Cuban socialism has been sustained for four reasons: improvement in material conditions; collective memories of making the Revolution and ongoing participation; ideology's stress on the benefits of the Revolution; and the convergence of a male gender ideal with that of the good revolutionary (p. 166). Indeed, Rosendahl suspects that the economic crisis of the 1990s might test the real and symbolic male leadership of the Revolution.

The treatment of power that centers the argument in *Inside the Revolution* is both prescient and problematic, as is the strong feminist critique that emerges. Rosendahl locates herself openly within the study, recounting expectations of repression, censorship, and more. If some fears were dispelled by the "much more open" society she discovered, other expectations are sustained in specific conclusions about the actualities of politics and power. Rosendahl owns the difficulties of "knowing what people 'really' think" (p. 5) and explains that over time she was able to move beyond a perception that most everyone was a "good revolutionary" to see "below the surface" (p. 21). What she sees there are indifference and varieties of nonparticipation that are all called resistance, protest, or opposition. She also finds "invisible power" (of social control) and understands power's "capriciousness" personally (through problems with offi-

cial permission for her research). Power is never explicitly defined, however, and Rosendahl assumes from the outset a monolithic official ideology that is at odds with the analytical possibilities present in her framework. Reinforcing this assumption, the either/or perspective on participation may overpoliticize ordinary people's behavior and, recalling that men and women differ little in their attitudes, depoliticize what is a compelling feminist reading of cultural and political practices—from food to internationalism. It is only in the epilogue's discussion of recent reforms that active negotiation expands to include adjustments to the official ideology, amending the centralist reading of its hegemony and freeing ordinary men and women from a powerlessness prefigured in the rich detail of Rosendahl's account of local political life.

Salt of the Earth: The Political Origins of Peasant Protest and Communist Revolution in China. By Ralph A. Thaxton, Jr. Berkeley and Los Angeles: University of California Press, 1997. Pp. xix+425. \$65.00.

R. David Arkush
University of Iowa

This book is innovative, interesting, and excessive in its claims. It is innovative because it is true oral history such as has rarely been done for China. Between 1985 and 1993, Ralph Thaxton conducted 1,000 hours of interviews with 200 villagers to piece together the histories of four north Chinese villages in the two decades leading up to the Communist Revolution of 1949. Although one wishes he showed more concern about the ways in which memories of events half a century earlier might be influenced by, for instance, years of official discourse on the evils of the "Kuomintang," still his narratives are basically convincing and show how oral history can enrich our understanding of the Chinese Revolution.

His informants tell an interesting story of a poor five-county area near the Yellow River, at the interstices of the three provinces of Hebei, Henan, and Shandong. Saline soil, flood, and drought made most families in his villages dependent for subsistence on producing and selling salt made from the earth. This business was illegal, for salt was a government monopoly and an important source of revenue for the central government. Much of Thaxton's attention is on conflicts between village salt makers and government revenue officers sent to destroy their salt ponds and confiscate their salt. Some of these conflicts involved attacks by thousands of armed farmers, and several succeeded in getting the revenue police withdrawn. Thaxton portrays the struggles as being against outsiders, not local elites. He carefully shows how Communists were not the original instigators but came to link up with and provide a degree of leadership for the movement. His evidence that villagers were seeking restoration of customary rights is less strong, but it is possible.

Thaxton begins to lose the reader's confidence, however, when he tries

to parlay this interesting local history of struggles over salt into a new interpretation of the Communist Revolution. The Communists, he claims, "ultimately won over the rural people by upholding their rights to produce homemade products for retail trade" (p. 258), and the "revolution was, to a significant extent, a broad popular struggle to preserve the market against . . . modern state making" (p. 319). His own data are not convincing. The salt struggles occurred in the early 1930s, which was not the period of Communist success in this area. Communist growth came later, during the Japanese occupation in 1937–45, but by then, salt was freely traded and no longer much at issue (the Japanese were shipping to Japan most of the government sea salt, produced around Tianjin, and had little interest or ability to stop village production or marketing). Thaxton's account seems more to confirm the standard explanation that villagers were drawn to the Communists by Japanese brutality and Communist fair, progressive taxation. Communist success continued in the 1946–48 civil war, a period when the Nationalists had no ability to enforce their salt monopoly in the countryside, and Thaxton's informants indicate it was due largely to land reform and the destructive Nationalist counterattack of 1947 (though he says a few spoke of fearing the return of the "Kuomintang salt police").

Contraband salt was obviously not an ordinary commodity, and it stretches credulity to equate business in it with marketing in general, calling government efforts against salt smugglers "Kuomintang antimarket attacks" and wartime Communist tolerance of salt business "defense, and even promotion, of the market" (p. 339). Seeing Nationalist revenue efforts as modern state building appears more reasonable, but one wishes Thaxton showed more awareness of the fact that the salt gabelle goes back two millennia in China and that conflicts between government revenue agents and local smugglers have been going on for centuries.

The book is also marred by sloppy and turgid writing, clichés, meaningless modifiers, mixed metaphors, invented words, and cumbersome constructions. "Puyang's Kuomintang power holders prudently cautioned of the potential incrimination from harmful interference in the celebration" (p. 132). Simple ideas are obfuscated by ponderous verbiage, which makes the book tedious and overlong. "Surely anti-Japanese patriotic appeals and rural socioeconomic reforms framed by the CCP weighed importantly in the development of Communist power during the war, but in much of the Hebei-Shandong-Henan border area the laurel of the CCP's involvement with the collective protest of the peasant saltmakers provides a major key to understanding the success of Communist-led resistance" (p. 199). He writes several times "land owned by X lineage" (see p. 169) when he seems to mean land owned by individual *families* in X lineage, not corporately owned by the lineage, an important distinction. Tables give no sources; one of "Peasant Earth Salt Proceeds in Monetary Units" (p. 298) has no monetary units and appears to be based on the author's imagination. The surnames of two well-known scholars, *Hsiao* Kung-chuan and *Chen* Yung-fa, are embarrassingly taken for given

names in the bibliography and index, and so on. Still and all, this at-times-irritating work adds much new and important detail to our knowledge of the Revolution in north China.

Movement Genesis: Social Movement Theory and the 1980s West German Peace Movement. By Steve Breyman. Boulder, Colo.: Westview Press, 1997. Pp. xvi+236. \$59.00.

David S. Meyer
CUNY/City College

In 1982, more than 400,000 people assembled in Bonn, the capital of the Federal Republic of Germany, to protest against the nuclear arms race in general and against NATO's deployment of Pershing II and ground-launched cruise missiles in particular. The demonstration, in a large park, was emblematic of modern social movements, with multiple platforms, numerous speakers, musicians, and people milling around, drifting from casual conversations to speech to song. Staged by a very broad coalition of groups, participants included moderate Social Democrats, fundamentalist environmentalist Greens, conventional Communists, pacifists, religious activists, conscientious objectors, and young people who wanted to be a part of the largest movement of their lives. It was an exciting time.

The FRG's chancellor, Helmut Schmidt, had asked NATO to prepare to deploy the so-called Euromissiles five years earlier as part of a strategy to link Europe more closely to the United States. In response to these missiles, seen as symptoms of new bellicosity, volatility, and danger in the arms race, peace movements grew up across Western Europe and the United States, perhaps nowhere more dramatically and powerfully than in West Germany. Within a year of the Bonn demonstration, the movement had precipitated a split in the governing coalition, leading to Chancellor Schmidt's resignation, the fall of the Social Democratic coalition government, and the beginning of the very long reign of Chancellor Helmut Kohl and the Christian Democrats. Still, Euromissiles were never deployed in Germany; under pressure from the peace movements, the United States and NATO adopted a new arms control posture that ultimately allowed the Soviet Union to take a new tack in its foreign and domestic policies. The Cold War ended, Germany unified, and the nuclear arms race will never be the same again.

The sketch above suggests how complicated and important this movement was. After the Cold War, the same movement offers a critical case to use in understanding the relationship of movements to the policy process, to more established and conventional means of political participation, and to the broader social and cultural trends from which they emerge. Steve Breyman, openly sympathetic to the West German activists, is less interested in providing a comprehensive narrative of the movement (for such an account, see Alice Holmes Cooper, *Paradoxes of Peace*:

German Peace Movements since 1945 [University of Michigan Press, 1995] than in using the case "to interrogate, criticize, and extend the theoretical literature on the etiology of social movements" (p. 1). Breyman is ecumenical in his theory, reviewing and making use of an impressive array of texts and approaches, examining the premises of six conflicting explanations identified as "party failure," "fear of war," "collective behavior," "resource mobilization," "political process," and "new social movement theory." He is committed to making use of all applicable insights, rather than accepting or rejecting any approach wholesale.

After providing sketches of each of the major theories he identifies, along with thumbnail critiques and criticisms, Breyman devotes the central portion of the book to substantive elements of the West German peace movement, addressing specifically the origins, organizations, goals, strategies, tactics, and mobilization of the movement. Relying mostly on published accounts in English and German, including some movement documents, Breyman identifies significant issues under each topic, and in doing so, sketches a profile of the movement. Recognizing the diversity of interests represented by the movement, he devotes a great deal of attention to the formal representation of groups within coalitions that coordinated movement activities, borrowing and offering theoretical underpinning to activists' ideas about movement spectrums.

Breyman's choice of a thematic rather than a narrative approach may be hard for readers not already familiar with the issues and events of the peace movements of the 1980s. The author's concerns here are, however, to apply the diverse theories of movements to the case, focusing particularly on the movement's origins. For those who know the basic outlines of the movement's trajectory, the book's chief strength is its openness to competing theoretical paradigms; this openness, however, also presents a challenge to the reader.

In his conclusion, Breyman uses his empirical material to examine particular aspects of each of the six theories he examined. Unsurprisingly, he finds that each of the theories has strengths and weaknesses in explaining this case. In the final few pages of the book, he tries to draw from this "amalgam of insights" to offer an "empirically-based theoretical account of the movement . . . that addresses the sources of social action: individual, organizational, communal, societal, transnational, global" (p. 197). It is impossible not to admire the theoretical ambitions of this work, but such a synthetic theory needs more than the allotted few paragraphs to be developed. *Movement Genesis* ends with what could be a beginning.

The Social Psychology of Protest. By Bert Klandermans. Oxford: Blackwell Publishers, 1997. Pp. ix+257. \$21.95 (paper).

James M. Jasper
New York City

Bert Klandermans has probably done more than anyone else to bridge American and European traditions of research on social movements. He addresses both approaches in his work, in which he has gone out and measured all sorts of things that no one else has. As important are the personal networks he has built through regular visits to the United States from his base in Amsterdam. As a sign of his influence here, he last year headed the American Sociological Association section on collective behavior and social movements.

As the title of his new book suggests, Klandermans also aspires to integrate the fields of social psychology and social movements by applying the former to the latter. In the first of two parts, he examines what draws individuals into participation, as well as the less studied issue of what causes them to end their participation. Drawing heavily but gracefully on his own (numerous) previously published articles, he integrates well-studied concepts such as the generation of collective action frames and the use of personal networks in recruitment with more rationalistic concepts such as costs and benefits and expectations of success. Klandermans's empirical cases deftly alternate with conceptual discussions. In part 2, concerning the organizational and political contexts of participation, the balance tilts toward reviewing the existing literature. Social movement organizations have been well studied (so much that they have their own acronym, SMOs), but less attention has been given to the fields of conflict within which they maneuver and seek out opportunities for action.

A thorough outline of the social psychology of protest would be a good idea and indeed Klandermans's title leads us to believe he will supply this. After all, the field of social movements has moved rapidly toward a social psychological perspective, usually under the strikingly similar banner of culture. Although these developments are directly relevant to most of his concerns, Klandermans takes little note of them, let alone integrate them into his models. He does not deliver on the promise of his title. For instance, in the first part of his book, Klandermans says he is examining individuals, but we meet only one individual protester, described in a brief paragraph taken from an obscure, unpublished work by another scholar. Instead we have aggregates of opinion polls, hardly the way I would go about getting at the social psychology of individual protesters. We have no in-depth interviews or life histories, no small-group dynamics, experimental evidence, psychological or emotional factors, cognitive biases in decision making, no moral sensibilities or principles. Assertions that individuals weigh costs and benefits are a way of avoiding social psychology, which offers the means to investigate why,

for complex cultural and psychological reasons, individuals value the costs and the benefits of protest as they do. Surveys are not the best means for getting at this. Klandermans cites a few classics of social psychology on information processing and communications reception, but otherwise uncovers little that would suggest that interesting processes are going on inside people's heads.

The Social Psychology of Protest, along with a recent book by Sidney Tarrow (*Power in Movement* [Cambridge University Press, 1994]), and an edited volume by Doug McAdam, John McCarthy, and Mayer Zald (*Comparative Perspectives on Social Movements* [Cambridge University Press, 1996]), feel like the summary chapters of a paradigm: the delimitation of a small and finite set of concepts for understanding the world, a disjuncture between the empirical evidence and the theoretical statements made about it, and "theoretical" works that are more literature review than creative theorizing. For 30 years, a relatively structural and organizational paradigm—once called the resource mobilization approach and more recently the political process approach—has inspired vast quantities of normal science by excluding a number of difficult phenomena. Many of these are knocking at the gate: cultural constructionism, cultural meanings and social psychology, individual biographies and psychologies, affective loyalties and emotional reactions, the psychology of strategic choice, the tactical, cultural, and organizational creativity of protesters, and much more. They have only occasionally been admitted, as in the use of "frame alignment" as a recruitment strategy. The old paradigm has too many important insights to vanish altogether, but its outer walls may be ready to come tumbling down. Bert Klandermans has written a fine scholarly book, which goes as far toward accommodating social psychology as one can go within the essentially structural root metaphor of an organizational paradigm. But it feels more like an end than a beginning.

Stalinism and Nazism: Dictatorships in Comparison. Edited by Ian Kershaw and Moshe Lewin. New York: Cambridge University Press, 1997. Pp. xii+369. \$54.95 (cloth); \$18.95 (paper).

Daniel Chirot
University of Washington

We all have opinions about Hitler and Stalin. Along with a handful of others, including Mao Zedong, they are the major examples of individuals whose influence may have transcended the social forces that produced them to permanently transform our world. Had Hitler's mixture of racial paranoia and arrogance been less extreme, would Germany have set out to exterminate all Jews? Would it have attacked the Soviet Union in 1941 and provoked the two-front war every German military leader had been taught to fear? Had Stalin not been so paranoid, would so many millions

have been enslaved, humiliated, tortured, and killed? Might a more rational, humane, and effective form of Communism have emerged, one that would have survived longer?

These are big questions that few sociologists have pursued. We study impersonal social forces, and if we look at history at all, it is usually to try to build transhistorical, decontextualized theoretical models. Nor do we often take ideology very seriously as a causal variable.

Stalinism and Nazism is an edited book that revisits old debates about Hitler and Stalin, discusses newer scholarship about their monstrous regimes, and compares them in light of some new Soviet documentation that has been uncovered since the late 1980s. It also reviews the bitter debates that have raged in German and Russian historiography about the causes and effects of the national calamities brought on by Communism and Nazism. The approach is refreshingly interdisciplinary and even includes some sociologists.

The general consensus is that the cruelty and chaotic inefficiency inside Nazi Germany was inherent to Nazism. (It has long been known by specialists, e.g., if not by the wider public, that Nazi Germany was less efficiently mobilized for war than the United Kingdom.) Hitler's system, according to the argument championed most famously by the German historian Hans Mommsen, was unstable and disintegrating from the start. It needed perpetual war to keep itself in power and could not have survived for long.

The book's consensus about Stalinism is quite different, thus explaining the title, which might otherwise have been either "Stalinism and Hitlerism" or "Communism and Nazism." It was Stalin's personality, his morbid paranoia, that was at the heart of the problem. Communism did not have to be so evil or ultimately self-destructive.

Is this consensus reasonable? Michael Mann's essay points out that Nazi Germany disintegrated because it lost the war, not vice versa. Its policy errors were driven by the logic of its revolutionary ideology, not by an inherent instability in its system. The Nazis were intent on changing the world; Hitler's visionary utopia of a racially purified Germany, Europe, and world dominated by a master race explains his choices. In a neat reversal of what we might have expected, the historian Mommsen's argument becomes a kind of sociological determinism based on organizational theory, while the sociologist Mann's point is that ideas mattered more than organization.

But none of the essays in the book are quite willing to recognize that revolutionary Marxist ideology was also inherently murderous, though on class rather than racial lines. To say this is the final taboo for anyone who still aspires to be on the left, however moderate. Yet there are ample comparative examples that show that many, if not quite all, Marxist regimes were driven by a kind of systemic paranoia that surely transcended any leader's personality. How could it have been otherwise? Dictatorships of tiny, embattled elites who thought it was their historical mission to overcome vastly powerful international and domestic class enemies

were unlikely to be gentle social democrats. Furthermore, one of the great appeals of Communism was that it promised to transform backward, previously dependent nations into powerful modern ones. But this meant either engaging in or preparing for war at all times. Communist societies were highly militarized. So, whatever revolutionary logic drove Nazism to its evil excesses was present in Communism, too, even if the theory was partly different. Stalin was not the only mass murderer and jailer of vast numbers of his compatriots among Communist leaders. Kim Il Sung, Enver Hoxha, Mao Zedong, Pol Pot, Haile Mengistu Meriam, and Sékou Touré may have been the worst of the others, but even such "benign" Communists as Tito (who killed tens of thousands before becoming a reformer) or Nicolae Ceausescu (who ruined and terrorized his country but killed rather few people) showed many of the same tendencies.

In the end, this book is a very useful, carefully done, sound, and provocative addition to the debates about the nature of revolutionary, ideological dictatorship, but it avoids some crucial questions that remain to be explored. Here, comparative historical sociology can make a significant contribution as long as it is willing to take ideology seriously as an independent variable and accepts as essential data the immense amounts of research already produced by the many historians working in this field.

The Policing of Politics in the Twentieth Century: Historical Perspectives. Edited by Mark Mazower. Providence, R.I.: Berghahn Books, 1997. Pp. vi+262. \$49.95.

Laura Kalmanowiecki
Yale University

State monopoly over the means of legitimate violence lies at the heart of the process of state formation. Professional policing is a creation of the modern state. The 19th century brought in Europe the civilianization of domestic politics: the police were entrusted with the maintenance of internal order, whereas the military were directed against foreign forces. Contrary to the myth of a consensual heyday for policing, scholars have shown that its origins should be explained as a response more to pressing political and social challenges than to long-standing problems of daily criminality. The threat—real or imagined—that successive governments confronted had a major impact on the style of policing adopted.

Rather than assuming—as Hsi-Huey Liang suggests (*The Rise of the Modern Police and the European State System* [Cambridge University Press, 1992])—the creation of police systems as a natural part of the modernization process, Mark Mazower has successfully located police systems at the heart of the organization of societies and state power. *The Policing of Politics in the Twentieth Century* constitutes an important contribution to the theoretization and systematization of the sociology of policing, which has come to age in the 1970s and 1980s. Some of its authors, such

as Cyrille Fijnaut, Clive Emsley, and Herbert Reinke, have previously collaborated at *Les Cahiers de la Sécurité Intérieure*, where they viewed the character of the police as a historical construct, centered around the state, and in relationship with politics, the legal system, and the demands of the policed.

In this volume, Mazower has emphasized the crucial importance of the rise of the modern state, acquiring an ever-increasing range of competences through the 20th century. It deals with the links between Europe and the rest of the world, and it illustrates the way different types of political regimes have come to terms with the need to establish devices to monitor subjects and civilians. Policing changed as a result of challenges to the existing distribution of power in society. Threats to established states and the fear of Communism after the Bolshevik Revolution led to the expansion of political policing, both in democratic and authoritarian regimes. The victory of Communism provided a permanent justification for the expansion of political policing. If counterespionage was at the center of much pre-1914 police activities, the fear of Communism prompted secret police to monitor "subversives." In Mazower's words, "The old nineteenth century distinction between opposition to government policies and treason crumbled, and with it an important defence protecting the individuals against the state" (p. 247).

As Emsley demonstrates, political policing in Europe constituted an instrument for states to provide information about different threats and helped to suppress them. Maintaining secrecy was justified as a functional need, and secrecy became the signature of secret police forces. After all, they were allegedly collecting "renseignements" for the "common good." A distinction was drawn between the British, more preventive style of policing and the more reactive one dominant in continental Europe. However, the benign nature of political policing did not prevail in Ireland or in the British colonies. As David Killingway shows, colonial rule could only be coercive since the aim of British colonial police officers was to maintain the *pax Britannica* (p. 169). Colonial policing was more concerned with internal security than with crime prevention. Furthermore, the compelling need to secure public order amid resistance, portrayed by Keith Jeffrey in Ireland, raises questions about the distinction between the police forces and the military for internal order maintenance. It could be argued that the functional distinction becomes blurred when their tasks overlap, as reflected by the military counterinsurgency role in Greece at the height of the civil war described by Mazower.

Mazower's work raises a pivotal and troubling issue: Is policing responsible for the pacification of societies? How do we measure the success or efficiency of political policing? Despite the Special Branch efforts, the British Empire vanished. Even more disturbing is the fact that, as Jean-Marc Berliere has shown, in France, political police often developed what it was fighting "by giving importance, means, attention and troops to groups that otherwise would have disappeared" (p. 47). Berliere demonstrates the constant temptation for the secret arm of the state to become

a state in itself, and for its agents to keep themselves in business by nourishing subversive threats that justify their own existence. Nor can we ignore the different ways in which practices that became routine parts of secret police repertoire, such as monitoring, surveillance, infiltration, shadowing, vetting, and covert operations (described by Athan Theoharis in the United States; Mazower in Greece; Berliere in France; Elise Tipton in Japan; Herbert Reinke in Germany), strongly affected the capacity of contestation and resistance of diverse actors.

Mazower is pessimistic about the prospects for shrinking security services in the absence of public concern. It should be noted that the persistence of secrecy and the autonomization of secret political policing conspire against public disclosure and accountability despite periodic oversight, as the recent FBI example illustrates.

In Mazower's view, "Bureaucratic empire-building, personal ambition and ideology explain at least as much and perhaps more than the existence of actual threats" (p. 254). We could add, however that institutional memories, past legacies, the demonstration effect of policing lessons, and mutual exchanges—including international cooperation in the definition and repression of threats (superbly portrayed by Fijnaut)—also shape policing regimes. This raises the troubling idea that, once institutionalized, the initial form, style, and organizational schemes adopted by policing regimes are quite resilient to change. Modern professional policing does not necessarily accompany political democracy, even after attempts at democratizing and professionalizing the police force.

Settling Accounts: Violence, Justice, and Accountability in Postsocialist Europe. By John Borneman. Princeton, N.J.: Princeton University Press, 1998. Pp. xii+197. \$49.50 (cloth); \$14.95 (paper).

Stanley Cohen
London School of Economics

Justice in transition, accountability, coming to terms with the past, democratization, truth commissions, reconciliation, "after the fall"—these are familiar terms in the public discourse about recent dramatic changes from authoritarian regimes toward democracy. The vexing question is how these societies deal with their dark pasts, especially the abuses of power, human rights violations, or crimes committed by the previous regime. Three very different clusters of cases have dominated the debate: first, the collapse of military juntas in various Latin American countries; second, the end of apartheid in South Africa; and third, the dismantling of Soviet and Eastern Bloc Communism.

John Borneman's study of accountability in the former East Germany (with some rather selective comparisons with other East-Central European states) is a valuable contribution to the expanding social scientific literature on transitional justice. Beyond his interest in this specific case

of decommunization, he has a more ambitious thesis about the centrality of the rule of law—more particularly, retributive justice—in legitimizing the emerging democratic order.

As an anthropologist, though, the most original and convincing parts of the book are his ethnographies of how taken-for-granted categories—criminality and the rule of law, perpetrator and victim, reconciliation and vindication—are politically constructed. Thus, under the heading “ethnography of criminality,” he shows how the *Rechtsstaat* was invoked through criminal justice. What did the new unified German judicial system deem to be “crime” in the former GDR? Borneman describes how this project was carried out by ZERV, the Central Investigative Office for Governmental and Unification Criminality, set up in 1991. “Unification criminality” covers economic crimes associated with unification, “governmental criminality” refers to crimes and human rights violations committed by party and state officials throughout the GDR history from 1949 to 1989. He shows in detail the role of evidence and archives (the 120 miles of shelves of documents stored in the Stasi Document Center) in what he nicely terms the “transformation from misfortune to injustice.” The consequent problems of accountability are illustrated by a well-known extortion trial from 1994.

Even more original is a section fascinatingly entitled “ethnography of vindication.” The complex theoretical links between retributive justice and restoration of dignity to the victim are grounded in a study of the workings of an East German Commission of Vindication/Rehabilitation. In certain firms and organizations, such commissions functioned from late 1989 through 1994 to restore the dignity of victims of the old regime through acts of vindication and rehabilitation. Borneman sees these—like the criminal trials—as staged public events, rituals to repair the damage done to the victim and to defeat the wrongdoer’s claim to mastery. They are performed in the belief that the performance itself will clarify what is right. His case study of the vindication proceedings for radio and television staff gives a rich sense of the micropolitics and personal lives behind the abstractions of “justice.”

As public events, though, these ceremonies of offender prosecution and victim vindication are, to say the least, more morally nuanced than these terms suggest. Two of Borneman’s many examples: first, “the most successful prosecutions have not been for typical or ‘normal’ forms of wrongdoing, but for excesses in the performance of public duties” (p. 143); second, former officials seek to deflect attention from their own complicity by reinventing themselves as *victims* of overzealous judicial reformers.

These empirical sections are framed by a general argument that I found less convincing. It is one matter to claim that accountability is central for emerging democracies and “is established in *part* [my emphasis] through retributive justice: a reckoning with the past where the government ritually purifies itself in periodic prosecutions of actual wrongdoing in the center” (p. 144). It is less evident that settling accounts by ritual criminal proceedings is absolutely necessary both to reestablish the dignity of vic-

tims and as a "key index" of state legitimacy. Nor does Borneman's rather uneven comparisons with other postsocialist states prove that without retributive justice, cycles of retributive justice will inevitably follow, directed either against an internal scapegoat or an external enemy. His typology of "jural restructuring" in East-Central Europe (radical regime change and some retributive justice; little regime change and little retributive justice; radical regime change and extensive retributive justice) is interesting. But it hardly shows the relationship between the use of retributive justice and the absence of subsequent cycles of violence, nor does his citation of the horrors in the former Yugoslavia. The case becomes even weaker if extended to transitions other than postcommunism.

Borneman notes that up to 1989, the German phrase for "overcoming/reckoning with the past" referred to the Nazi past—and proved "impossible to fix." His thought-provoking study shows how the unified government's mandate of "overcoming of/reckoning with the GDR's past through criminal law" is just as troublesome.

The Great Surprise of the Small Transformation: The Demise of Communism and the Rise of the Private Sector in Hungary. By Ákos Róna-Tas. Ann Arbor: University of Michigan Press, 1997. Pp. xv+289. \$49.50.

Tony Levitas
RTI

Ákos Róna-Tas's *Great Surprise of the Small Transformation* represents one of the first sustained attempts to bridge the intellectual gap between the study of communism per se and the study of what has happened in its wake. Róna-Tas attempts to bridge this gap by linking the answers to two important and apparently disparate questions: Why did communism collapse? And what forces account for the particular trajectories privatization has taken in postcommunist countries?

In the spirit of Polanyi, Róna-Tas focuses on the labor market in answering the first question. He begins by elegantly recapitulating arguments that identify the essence of Marxist-Leninism in the effort to organize society as if it were a single factory. And he shows how in practice this entailed making the state both a universal owner and a universal employer. Moreover, he argues convincingly that by the 1960s, and after the CP itself, full employment had become the central institution for political legitimization and control under communism.

At the same time, he claims that in order to compensate for the shortages engendered by organizing society as if it were a single factory, the party began to loosen restrictions on the private sector, particularly in the production of consumer goods. This is fair enough, though Róna-Tas seems a little uncertain at times about whether he wants to argue that the economic problems of communism were caused primarily by the elimination of competitive labor markets or by a larger set of forces. The

problem for him here is that the former position is no doubt too reductionist while the latter position weakens the—though maybe less than he may have feared—linkage he is trying to draw between the forces that led to communism's collapse and the trajectory privatization took in post-communist Hungary.

In any case, he continues his argument by showing how the centrifugal tendencies unleashed by the partial accommodation of a rump private sector had begun to effect that state sector by the 1960s, leading to greater autonomy for industrial firms and their managers. Indeed, a hardening of the party line in the 1970s failed to fully stop the deconcentration of economic power that had begun earlier. But the real break happened in the early 1980s when party officials began to see the private sector and quasi-competitive labor markets not simply as a safety valve for the production of consumer goods but as forces that could be used to revitalize socialism.

In this, they proved misguided. In fact, throughout the book, Róna-Tas provides no truly compelling argument for why the Hungarian Party first started on this track and then, more or less continuously, held to it. Nonetheless, he does demonstrate that the shift in party policy allowed for both the rapid expansion of the private sector *per se* and the equally important proliferation of quasi-private activities within the state sector itself. The proliferation of these new entrepreneurial structures, in turn, had profound consequence for the nature of communism's collapse in Hungary, as well as on the strategies of privatization that were pursued after the fall: First, Róna-Tas argues—with compelling and well-marshaled data—that the communist elite had become so profitably engaged in the private sector that they could afford to give up political power without too much of a fight. Second, and more interestingly, he argues that the blurring of the boundaries between the private and state sectors that had taken place in Hungary during the 1980s made voucher-type privatization schemes unpopular. In short, key insiders—both managers and skilled workers—proved strong enough to block all privatization schemes that would have taken control over the pace and nature of the process out of their hands. Here Róna-Tas is clearly on very strong ground. Moreover, the general story he tells is well told, well documented, and will no doubt be of use in the field for years to come.

At the same time, it is a shame that Róna-Tas failed to make use of his mentor's concept of a "double movement": Polanyi, it will be remembered, used this concept to describe the commodification of labor that took place during the 19th century as well as the countermovements that immediately arose to limit its full impact—countermovements that he argued ultimately resulted in the emergence of the welfare state. Something similar seems to have happened under Soviet socialism, with the strength and effects of the double movement being less a product of the system itself than in the nature of its external imposition and maintenance. Indeed, the most interesting question may be what organizational forms might have developed out of communism's misguided attempt to

organize society as a single factory if its "double movement" had not been constrained by Soviet power.

The Testimony of Lives: Narrative and Memory in Post-Soviet Latvia. By Vieda Skultans. New York: Routledge, 1997. Pp. xxi+217. \$75.00 (cloth); \$24.99 (paper).

Mikk Titma
Stanford University

The opening to the world of the lands formerly in the Soviet bloc is a major historical event, leading to extensive scholarship on how society really functioned under Soviet rule. *The Testimony of Lives* by V. Skultans typifies an aspect of such work. It aims for depth on a narrow topic, not a broad picture. Through personal letters and oral narratives collected during visits to Latvia in the 1990s, Skultans, an anthropologist trained in England, chooses to rediscover her Latvian homeland, which her family left during World War II.

The Testimony of Lives has three main parts. The first (chaps. 1–3) relates the author's discovery of her homeland, her family history, and her research approach. The second (chaps. 4–9) contains her analyses of oral narratives she collected from a few dozen people. The third (chap. 10 and the appendix) describes Latvia and gives her overview of its history.

Latvia is a small European country on the southern coast of the Baltic Sea in the historical "no-man's land" between the German and Russian nations. Latvians as a nation were actors only during World Wars I and II. World War I brought suffering but also the first state belonging to Latvians. World War II took statehood away from Latvians and placed it in the Soviet domain. Most personal histories analyzed by Skultans reflect on this second national experience.

There is an ongoing debate about the subjectivity of perception of reality, especially of history in different cultural contexts; a nation develops its own view of history. For example, in Latvian history, the 16th-century invasion of Latvia by Russian Tsar Ivan IV ("The Terrible") is seen as a massive slaughter of people. In Russian history, Ivan IV is seen as unifying all Russian territories. Interpretations of the same events inevitably vary across national contexts, and these varying interpretations form a key element of a nation's unique heritage.

Skultans describes how she collected and documented oral narratives to lay the basis for a written version of the national memory of Latvia in the mid-20th century. "My approach does not challenge the truth of the past as witnessed, but rather investigates the cultural resources used to make sense of the past and incorporate it into a personal history" (p. 27).

The destiny of the Latvian people appears to be the broader cultural context into which Skultans positions personal histories. This worldview

is typical in small nations. The nation's existence demands a constant struggle for survival among more powerful neighbors and forms the backdrop against which its members evaluate their lives. While the gaining of Latvia's independence after World War I heightened the image of a national destiny and gave an enlarged cultural context for individual life, World War II brought misery and the loss of a Latvian state. It shook an important foundation upon which individual life had meaning and destiny. The cruel destiny of the Latvian nation cast its imprint on the individual lives described in Skultans' narratives.

Her chapter on "The Expropriation of Biography" points to another major aspect of the narratives. Every conqueror of Latvia rewrote its history, and the Soviets were no exception. But no previous conqueror forced people to change their personal and even their family history to survive and succeed. This experience is unique. Some people interviewed by Skultans requested anonymity because they are still in a vulnerable position.

A special chapter discusses the "forest brothers," as post-World War II guerilla fighters were called in the Baltic. The Soviets managed to subdue them in the Baltic region and western Ukraine only through mass deportation in March 1949. This deportation crushed all thoughts of resistance and made personal survival the primary issue. In the new psychological climate, fear was a *constant* element of personal life.

Chapters 8 and 9 are on the meaning of life. "Latvian narrators are preoccupied with meaning in a quite explicit way: they complain that their lives lack meaning and purpose" (p. 124). In a nation visibly losing ground and a society with a limited range of personal goals, memories lack a thread giving meaning to a harsh personal life history. It is hard for the objective analyst to know to what extent the meaning of personal life was destroyed by the communist system per se and to what extent by the unhappy fate of the Latvian nation in the Soviet Union.

For American readers who are not specialists, *The Testimony of Lives* may seem to be mainly a story about a small nation and its people. Its extensive details and interpretations of the meaning of Latvian words are sometimes tedious, but this is how anthropological treatises are often written—they are usually intended for specialists and not a large audience.

Power and the Sacred in Revolutionary Russia: Religious Activists in the Village. By Glennys Young. University Park: Pennsylvania State University Press, 1997. Pp. xvii+307. \$47.50.

Richard L. Hernandez
Stanford University

In this pioneering study of rural religion and politics under the Bolshevik regime, Glennys Young goes far toward filling in one of the least noted

but most deleterious "blank spots" in contemporary historiography of the Soviet Union. Until recently, the field has implicitly suggested through its own silence that religious belief during this period was in quiet decline or that it simply became the passive object of state-sponsored persecution. While both of these misconceptions are based on the reality of Bolshevik attempts to eradicate religion by the hook of social engineering or the crook of brutal repression, they nevertheless leave us ignorant of the everyday plight of believers under this regime. Young, however, succeeds at restoring a credible sense of the historical agency available to religious believers in the face of, indeed often in response to, the state's demands. Aside from a brief but thorough introduction to the rural religious scene at the end of the tsarist era, the author focuses mainly on the first decade of Bolshevik rule, at which time a cultural conflict raged over what role religion was to play in the new village. In light of the aforementioned scholarly neglect and the persistent assumption that Russian peasants inhabited an apolitical world, Young's conclusion is surprising. During the 1920s, religion remained so important to peasants' social identity that it mobilized rural clergy and laity to "revive factional politics in the Soviet countryside" (p. 273)—a state of affairs that eventually helped motivate the regime's brutal war on the peasantry during collectivization in the early 1930s.

Chapters 1 and 2 search out the historical beginnings of clerical and lay religious activism. Here, in a masterful synthesis of previous scholarship supplemented by her own research, Young locates the antecedents of religious politics in the tsarist "Great Reforms" era and in the first few years after the October Revolution of 1917, at which time rural clergy and laity faced the increasing pressures of social modernization as well as the political and cultural designs of the state. Chapters 3 and 4 survey the impact of official antireligious efforts in the countryside, suggesting that aspirations toward a godless utopia were largely frustrated by religious believers' indefatigable political activity. The book's remaining five chapters eschew chronological organization in favor of thematic analysis—covering the relationship between clergy and laity, the laicization of parish life, the co-option of Soviet political institutions by religious believers, and the vicissitudes of Bolshevik antireligious rhetorical categories.

To her credit, Young is not wedded to any single set of historiographical questions or methods but approaches her topic from a variety of angles in order to tell a complex, sometimes paradoxical, story. Social history has a prominent place here as the author highlights generational conflict, social mobility, gender, and economic conditions as important factors in the rural religious conflict. Explorations of cultural issues, however, produce the book's most important insights. Treatments of law, legality, and administration within the village's religiopolitical milieu, for instance, are valuable contributions to our growing knowledge of the accompanying complexities and ironies of Soviet modernity in the countryside. The 1918 Bolshevik "Decree on the Separation of Church and State," for example, was in many ways quite successful as a secularizing

instrument as it delegitimized the church's traditional structures and disenfranchised its hierarchy. Yet, because the decree allowed limited religious activity if strict legal requirements were met, it also gave religious believers in the village a modicum of juridical legitimacy and a concomitant rhetorical means to defend themselves against the regime's more militant antireligious activists. Moreover, by compelling interaction with the state's agents (through petition drives, tax payments, approvals for church upkeep, and so on), legislative or administrative efforts to control or repress ended up baptizing peasants into the world of modern politics and political technique. This is especially evident in electoral politics through which religious believers managed to "infiltrate" (a highly charged Bolshevik term) local Soviet administrative organs.

The book also makes important, though sometimes limited, contributions to our understanding of what Young labels "village semiotics." There are only brief discussions of apocalyptic discourse and religious images (icons) here, but there are also sustained and extremely insightful explorations of how the village church served as a crucial focal point of cultural struggle. Its enduring significance as a symbol for so much of what the regime attempted to destroy in the village—namely a politically independent peasantry with a competing cultural identity—explains why the activity of religious and antireligious activists alike so often centered on the sacred space of the church and why the destruction of churches became a paramount concern for the regime in the decades that followed.

The book is written in a clear style, never overburdened by theoretical language. At times, however, it would have benefited from a more rigorous theoretical defense. For example, the relationship between religion, or culture in general, and power appears too unidirectional in places where the functionalist biases of the sources (nearly all Bolshevik) are not always checked by the author's own interpretive framework. This minor criticism aside, Young has written a crucial and seminal book that will, I hope, spur a new wave of studies of religious identity and politics in the former Soviet Union.

Sacred Tensions: Modernity and Religious Transformation in Malaysia. By Raymond L. M. Lee and Susan E. Ackerman. Columbia: University of South Carolina Press, 1997. Pp. xi+172. \$29.95.

Robert W. Hefner
Boston University

Few countries present a more complex religious profile than the Southeast Asian nation of Malaysia. Ethnic Malays are officially recognized as the country's indigenous "sons of the soil" (*bumiputera*) and constitute 56% of the population. Malays are Muslim, and Islam is the national religion. Descendants of 19th-century traders and laborers, Chinese compose 32% of the population and are predominantly, if heterogeneously, Buddhist.

Indians, at 8% of the population, profess widely varying versions of Hinduism. Besides these primary groupings, Malaysia also has Sikhs, Baha'is, tribal animists, and a small but growing Christian community.

In this short, densely argued book, Raymond Lee and Susan Ackerman set out to make sense of this pluralist cacophony, explaining its origins in the early modern era and the dynamics of the religious revitalization sweeping Malaysia today. Though the authors have conducted extensive research on Malaysian religions, the book's argument is presented not as a summary of primary research but as an overview and reanalysis of literature on religion and modernity in Malaysia as a whole. Though at times this makes it difficult to assess the empirical grounds for their generalizations, the approach allows the authors to stand back and compare developments across Malaysia's ethnoreligious mosaic.

The authors begin with a discussion of Islam and the Malaysian state. They show that, prior to the colonial period, Malay Islam was raja-centered and little concerned with a detailed implementation of Islamic law (*shari'a*). The arrival of British colonialism in the late 19th century, however, coincided with the ascent of religious reformism in the Muslim world. Through their efforts to bureaucratize the administration of Islamic law, the British unwittingly undermined the religious authority of Muslim rulers and strengthened the hand of religious scholars. Equally important, the bureaucracy provided "an organized means for the definition of Malayness" (p. 33).

Since the 1960s, the authors reveal, an otherwise secular Malay elite has seen fit to promote a "highly Protestantized" (p. 36) form of Islam so as to strengthen Malay unity and inculcate a work ethic consonant with state-promoted capitalism. The elite failed to anticipate that some Malays would put the new orthodoxy to oppositional use, invoking Islam to voice their dissatisfaction with industrialization and its bureaucratic patrons. Challenged in the 1970s by a fundamentalist upsurge, the state bid up its commitment to Islam, obliging ordinary Malays to conform to a more unitary expression of their faith.

The situation of Malaysia's Buddhists and Hindus, Lee and Ackerman show, differs greatly from that of the Malays. Deeply involved in the institutions of colonial capitalism and education, the Chinese community was internally diverse from the start. Though resinification has taken place since 1969, many in the Chinese community are English-speaking and internationalist in their cultural and economic orientation. Responding to the Islamic resurgence among Malays, Chinese Malaysians sought to deepen their understanding of Buddhism. Many English-speaking Chinese did so by looking to Western Buddhism for inspiration, preferring its mediation and intellectualized study to the ritual and asceticism of its Asian counterparts. Even in the aftermath of efforts by Buddhist reformists to mobilize the laity and demystify folk Buddhism, however, Malaysian Buddhism remains heterogeneous and ecumenically nonsectarian.

Though considerably smaller, the Indian-Hindu population also shows

considerable internal diversity. The paucity of Brahmins among this immigrant population made caste concerns less divisive than in India. But caste divisions have been maintained nonetheless through marriage arrangements, occupational differentiation, and ritual roles. Brought by low-caste Tamils, popular Hindu worship focused on village and guardian spirits rather than Sanskritic deities. Over time, some in the Hindu community have sought to elevate the status of their temple deities through "grander temple reconstruction and ritual refurbishment" (p. 96). Others, however, have experimented with more direct forms of worship such as devotional hymns deemphasizing priestly ritual roles or orgiastic rites of penance and possession. Though the Islamic revival among Malays has given rise calls for Hindu unity, the community remains culturally and organizationally fissiparous.

Sacred Tensions provides an original and insightful summary of religious change in Malaysia's varied religious communities. Some readers will find the authors' strict constructionist Weberian framework, with its broad generalizations as to the civilizational concerns of the "Western mind" (p. x), one of the book's less attractive features. Even here, however, the authors take pains to point out Asia is not replicating the Western history of secularization, and their comments on the way in which global cultural transfers are affecting Malaysian religions are convincing. For courses on Southeast Asian religions or the general sociologist of religion interested in an intriguing non-Western case study, then, this is a useful and welcome book.

Millennium, Messiahs, and Mayhem: Contemporary Apocalyptic Movements. Edited by Thomas Robbins and Susan J. Palmer. New York: Routledge, 1997. Pp. ix+334. \$18.95 (paper).

James T. Richardson
University of Nevada, Reno

This is a fine collection of high-quality papers concerning millenarianism in Western culture. Some may think the effort too ambitious and the coverage too varied. But others will be appreciative of the attempt to relate scholarship in several areas that are sometimes thought of as independent of each other. The collection is obviously designed to speak to the growing interest in millenarianism among scholars and the general public as we approach that magic date of 2000. The volume will serve well as a text for special courses on that phenomenon, as well as for sociology of religion and social movements' courses.

The volume is divided somewhat arbitrarily into four major sections: the first focuses on "theories of apocalypticism," the second contains papers on "secularizing the millennium," the third is called "apocalypticism and the churches," and the fourth covers the important issue of "violence and confrontation." Major scholars in the sociology of religion, social

movements, and religious studies are contributors, including such figures as Tom Robbins, Dick Anthony, Michael Barkum, John Hall, Catherine Wessinger, Maasimo Introvigne, Mark Mullins, Anson Shupe, Jim Aho, Robert Balch, and David Bromley.

The theory section is strong, with a lead paper by Bromley that offers his usual well-grounded theoretical analysis of apocalyptic themes in Western culture. He proposes a structural as opposed to a theological analysis as being more valuable for understanding such views. The second paper by Wessinger is short and to the point but will be much-cited for its commonsense proposal that scholars drop certain mystical terms such as pre- and post-millenarianism and instead use her more general concepts of catastrophic and progressive millennialism.

Jim Aho offers a thoughtful essay relating postmodernism and millenarianism, proposing the concept of the "Linguistic Turn," by which he means the evolution of reality into a linguistic construct. This perspective is sharply contrasted to the rise of fundamentalism, with its millenarian thrust. Aho's essay contains some of the most sobering lines in the volume: "The Linguistic Turn calls us to step, eyes open, into the emptiness left after Man's disappearance. Fundamentalism urges us to flee the emptiness, to fend off its harbingers, to build impregnable fortresses that might keep chaos at bay (pp. 65-66). "To a handful of academics at elite Euro-American universities the end of Man means the rise of words. To millions, if not billions, of others man's erasure signals the resurrection of the Word, the rebirth of a transcendent center—this, either in the form of a personal deity or an impersonal force" (p. 68).

The theory section closes with a nice piece of empirical research by Rob Balch and a team of students, focusing on what happens with the repeated failure of prophecies. The research thus goes beyond the earlier classical work of Festinger et al. on the failure of prophecy, uncovering important distinctions between the reactions of leaders and followers and also offering a detailed analysis of the types of accounts and justifications that develop in the face of repeated failed prophecies.

The next section was quite enlightening, with its focus on more secular versions of millenarian thinking. Some will take issue with the thrust of Philip Lany's piece of the secular millenarianism of militia groups, but the details offered in the analysis are valuable. Martha Lees's paper outlining the ideology of 'Earth First!' is quite useful, helping us understand the motivations of individuals who established this controversial movement, as well as the pervasive tensions that developed within that group between those promoting "biocentrism" and those identifying with "biocentric equality." The former perspective retains a prominent place for humanity, whereas the latter view assumes that humanity is just one species among many, with no special status—a difference of perspective with profound implications and a difference that destroyed Earth First! as an effective organization.

John Bozeman's paper on "technological millenarianism" offers a fascinating interpretation of many science-based movements such as eugenics,

cryonics, interest in establishing space colonies, and similar ideas in recent American history. His positing of technological millenarianism with American culture will be off-putting to some, while prescient to others. Palmer's chapter on women in new religions is fascinating but may seem oddly placed in this section. Her focus is, however, on the feminization of such movements, or, better put, on the unrecognized role of females in apocalyptic movements in American history. The detailed information offered about the role of women in many contemporary "new religions" makes this a valuable chapter indeed.

The section on apocalypticism in the churches opens with Michael Cuneo's valuable study of Catholic apocalypticism, featuring discussions of the Fatima Crusade, the controversial Apostles of Infinite Love, and Veronica Lueken's Bayside Movement. Shupe's delineation of "Christian Reconstructionism" within the Christian Right is also enlightening, as well as disturbing. Ron Lawson's paper on tensions caused by recurring apocalypticism within Seventh-Day Adventism is informative, as are the details given about strong denominational tendencies within Adventism that move the movement beyond overt apocalypticism. The section closes with an excellent piece by the European scholar Introvigne on millenarianism within contemporary Mormonism. Introvigne notes the tension deriving from the inherent millenarian message of Mormonism and describes the many forms this tension takes as it is acted out by various segments and individuals within the Mormon Church.

The last section, on violence and millenarianism, opens with an essay by noted scholar Michael Barkun on the elements of thought of the Christian Identity movement, a movement that has spawned violence in recent history. Barkun's application of Roy Wallis's deviance-amplification theory to Christian Identity is insightful, and he points out possible limitations of that theory when apocalyptic beliefs are held as strongly as they are with Identity participants, especially in situations such as the U.S. context with its Constitutional protection for religious beliefs. Robbins and Anthony's chapter will concern some because of its implication of the inevitability of violence in certain types of religious movements, as well as its psychodynamic perspective. They develop a theory derived from their own work and that of Erik Erikson to explain why certain groups might become violent and apply that theory to the Waco tragedy, offering some valuable but disquieting insights.

John Hall and Philip Schuyler examine the bizarre deaths of Solar temple members in Europe and Canada, using a complex theoretical approach that focuses on three issues: (1) the formula for "salvation" of the predominantly Catholic group, which was drawn from selected counter-cultural elements; (2) struggles for cultural legitimacy of the group, which were ultimately unsuccessful; and (3) tensions between views of reality among members and leaders of the Solar Temple. Their lengthy analysis is quite informative and compelling.

The book closes with Mullins's informative analysis of the Aum Shinrikyo movement that brought such terror to Japan in 1995, with its deadly

attack using sarin gas on the Tokyo subway. This new Japanese religion is basically Buddhist in orientation but with a strong overlay of apocalyptic beliefs built on the idea of the impending nuclear self-destruction of humanity. Its organizational structure is deliberately modeled after that of the Japanese government, which suggests a very different sort of religious movement indeed. The violence that occurred within the movement prior to the gas episodes and the episodes themselves leave no room to avoid the conclusion that religious groups can "go bad" and cause much harm to members and others.

The entire volume is a solid contribution and raises many issues for scholars, citizens, and governments. Among them is the perceptive closing statement of Mullins's paper: "The difficult task ahead is to find a way to preserve the free practice of religion in modern societies, while at the same time keeping dangerous materials and technology out of the hands of those who would use it to bring Armageddon and establish their particular version of the millennium" (p. 322).

Contemporary American Religion: An Ethnographic Reader. Edited by Penny Edgell Becker and Nancy L. Eiesland. Walnut Creek, Calif.: AltaMira Press, 1997. Pp. 272. \$46.00 (cloth); \$22.95 (paper).

Daniel V. A. Olson
Indiana University, South Bend

During the 1960s and 1970s, many of the advances in sociology of religion came through statistical analyses of broad patterns apparent in survey research of individuals and congregations. More recently, spurred by generous support from the Lilly Endowment and the Pew Charitable Trust, the sociology of religion has experienced a boom of ethnographies of seminaries, congregations, and small spiritually based groups.

This book's main strength is that its chapters represent some of the best of these ethnographies. All are based on quite thorough research, some lasting for several years, and one involving as many as 230 interviews. The researchers draw on a variety of sources: personal interviews, participant observation, census data, surveys of members, and organizational records. They range in scope from an intensive examination of the interwoven religious life histories of several members of a single, alcohol-plagued family to an examination of conflicts within 23 congregations in a single suburban community. All are clearly written, and each seeks in some way to debunk, qualify, or expand previous generalizations about American religious behavior.

A secondary strength is that individually the chapters offer new information and many interesting insights about contemporary American religion. Several chapters explore the ways in which identity is constructed, maintained, or abandoned among persons trying to combine identities that are often seen as opposed (e.g., Messianic [Christian] Jews and femi-

nist evangelical women). Similarly, another chapter explores the serious religious significance and ambiguous identity dynamics behind the long running, quite popular performances of the "Gospel Hour," a weekly Sunday "service" of traditional gospel hymns sung to a "congregation" of gay men in an Atlanta bar by the Gospel Girls, a straight woman and two gay men in drag.

Other chapters cover a range of issues: styles of moral reasoning found in two, quite different, pro-life groups; the distrust found among residents of a poor, crime ridden, inner-city neighborhood and the ways in which this distrust affects, and limits, relationships in a local African-American congregation; the ways in which an Atlanta Baha'i community locally enacts its global ideology; how smaller congregations respond to the growth of a nearby megachurch; and the way arguments are framed during conflicts in local congregations. This last study, by one of the editors, shows that many of the issues causing conflict in the 23 congregations she examines are "culture war" issues (e.g., sexual behavior and identity, women's roles, race). However, contrary to James Davison Hunter's depiction of how these issues are debated in the public media, congregants are driven not only by the need to determine what is morally right but also by a religious concern that the solutions be caring and respectful for all concerned. This second dynamic undercuts much of what some see as an inevitable tendency for religiously based conflicts to become uncompromising, winner-take-all fights to the finish.

The concluding chapter by Robert Wuthnow, one of the persons most responsible for the increase in ethnographies of American religion, highlights the major insights of each chapter and offers a balanced perspective on the contribution of ethnographies to the study of religion. He appreciates that ethnographies have put a renewed and rightful emphasis on the cultural aspects of religion but warns against a one-sided reliance on ethnography to the exclusion of other research methods.

As the above description of topics suggests, the book's chief weakness, especially from the point of view of textbook sales, is that the chapters have few common themes. Individually, the chapters touch on a broad range of issues debated in sociology of religion: secularization, the role of strictness, competition in religion, culture wars, and the role of religion in the dynamics of identity formation. However, with the exception of identity issues, there is little sustained analysis of a single topic across more than one article. Nor should readers expect a summary chapter showing how each of the studies point to some grand picture of changes in American religion. In fact, the editors suggest that a major advantage of ethnography is that it debunks grand theories and simple explanations.

Though lack of an explicit common theme is a weakness, the book's chapters, taken together, highlight an implicit theme: the incredible diversity of ways in which Americans constantly create and recreate meaning, belonging, and personal identity in and through religious culture. Religious ideas, identities, and practices are not merely created by distant religious producers (elites) for local sale to consumers who chose to pur-

chase one model of religion versus another. Rather, American religious culture is constantly being remade and refashioned locally by the supposed consumers, sometimes in quite unusual and unexpected ways.

Cultural Wars in American Politics: Critical Reviews of a Popular Myth. Edited by Rhys H. Williams. New York: Aldine De Gruyter, 1997. Pp. xii+299.

Wade Clark Roof
University of California, Santa Barbara

Since the early nineties, the metaphor of a cultural war in the United States has attracted considerable attention among pundits, preachers, politicians, and professional social scientists. The metaphor, associated particularly with sociologist James Davison Hunter, resonates in a climate where values, morality, and lifestyles are contested and at a time when, for social scientists, debates over old cleavages based upon class, race, region, and other social-structural differences seem to have lost luster. But increasingly, there is critical attention to this cultural wars thesis and to what Rhys H. Williams, editor of this volume, regards as "overheated rhetoric." Assembled in the volume are 12 rather substantial essays: some new, some previously published, all directly addressing the thesis. Plus it contains a thoughtful afterword by the editor.

Many of the essays draw upon national survey data, mostly the General Social Surveys, for mapping the contemporary cultural terrain. The contributors are N. J. Demerath III and Yonghe Yang; Nancy J. Davis and Robert V. Robinson; Paul Dimaggio, John Evans, and Bethany Bryson; Ted G. Jellen; and Daniel V. A. Olson. No matter what the substantive focus or the measurement and statistical procedures employed, the results are pretty much the same: while there is some support for a notion of cultural conflict along an orthodox versus progressive continuum, that evidence is strongest at the ideological extremes and fails to fall out along a single axis. A consistent finding is that the data form at least two axes, one pertaining to economic and political matters and another with regard to personal moral and lifestyle issues. Often the two dimensions are orthogonal, which, as Williams concludes, is basis for at least a fourfold distinction rather than a simple binary conception of a cultural divide. In this respect, more recent American opinion seems structured not all that differently now than in the past, leading one to question what is really new, if anything, in the so-called culture wars except the particular issues and how they became framed in the eighties and nineties.

Other essays examine elites, churchgoers, seminaries, political elections, intergroup conflict, and the United States in comparison with other countries. Timothy Shortell finds that strong religious commitment is not always associated with greater conservatism and that, on economic indicators, liberal Protestants can be as conservative as any others. Christian

Smith and associates observe that many Americans, including conservatives, in making moral judgments rely upon both a transcendent authority as well as the authority of experience, which often neutralizes social activism. Jackson W. Carroll and Penny Long Marler look at two Protestant seminaries, one liberal, one conservative, and conclude that, while there are warring moral visions in American culture grounded in religious tradition, the more intense battles are often not between traditions but internal to them. Margaret S. Hrezo and Melinda Bollar Wagner examine Oliver North's campaign for the U.S. Senate in Virginia, observing that "realism" prevailed over "idealism" and that for the Christian Right to be successful in American politics it would have to accommodate the cultural mainstream. Gerald M. Platt and Rhys H. Williams distinguish between "individualist" and "collectivist" bases of ideological divide in American politics yet caution against overinterpretation of that divide. Fred Kniss proposes an alternative, two-dimensional model of intergroup conflicts, distinguishing between the locus of moral authority and the "moral project," whether the individual or the collectivity. N. J. Demerath III and Karen S. Straight, using Demerath's cross-national observations, offer five analytic distinctions in relations of states and religions and argue it is unlikely for the United States to have large-scale cultural violence over state power. These essays underscore American religious and political institutions as complex, crisscrossing realities that impede any thoroughgoing polarization.

Williams's afterword is especially helpful not only in summarizing what is learned from the essays but in laying out the analytic issues that must be attended to in addressing what amounts to a popular myth of a great American cultural divide. Aside from getting the data properly ordered representing the American populace, it is essential to distinguish elite rhetoric from broadly based opinion; to recognize that the bases on which people form their opinions and sentiments tend to be unstable, depending upon the issues and levels of salience; to understand that connections between worldview, opinion, and action are constructed in particular social contexts and do not necessarily follow in a predictable order; and to sort out the differing logics for centrist-driven institutions and movement politics, which inflate identity and ideological differences. Any social scientist hoping to move the discussion on moral, religious, and political divisions in the country forward by relying less upon rhetoric will find the volume useful.

Making American Industry Safe for Democracy: Comparative Perspectives on the State and Employee Representation in the Era of World War I. By Jeffrey Haydu. Urbana: University of Illinois Press, 1997. Pp. x+261. \$49.95 (cloth); \$21.95 (paper).

Kathleen Thelen
Northwestern University

Jeffrey Haydu's *Making American Industry Safe for Democracy* is a well-conceived and beautifully executed comparative analysis of the impact of state policy on industrial relations during and after World War I. Although the main focus is on the United States, the author employs brief but effective comparisons to Germany and Britain in order to highlight how political and industrial conditions shape state policy. The book's central argument is that state policy to resolve the wartime labor crisis contributed significantly to the postwar industrial order that emerged.

The United States, Britain, and Germany all faced an upsurge in wartime militancy, but central to Haydu's analysis is the idea that these countries in fact faced rather *different* crises, which varied as a result of prewar labor conditions. What distinguished Germany from both the United States and Britain is that, in the former, labor militancy posed a threat to the *political* order, and so state policy addressed the political dimension by imposing a statutory solution that restored order by incorporating moderate elements of the union movement into new institutional arrangements. The comparison between Britain and the United States—both cases in which labor mobilization was less a political threat than it was an industrial relations crisis—highlights the importance of preexisting class relations to the type of crisis each state confronted as well as the solutions government elites sought. In Britain, unions were well established before the war, and so state policy sought to reimpose order by shoring up the authority of central unions against rank and file militants (in that sense similar to Germany, although British policy encouraged voluntaristic not statutory institutions). In the United States by contrast, unions entered the war in a much more tenuous position. Here employers coded unions, no matter how moderate, as part of the problem rather than part of the solution. In this case, then, government policy toward labor during the war preserved the open shop even as it sought to introduce and disseminate "enlightened management" practices that would restore order. In sum, whereas, in both Germany and Britain, government policy aimed to shore up central union authority and insulate union leaders from militant shop-floor movements, in the United States, government policy, if anything, helped sever the link between central union leaders and shop-floor based organizations, in the process, weakening unions in the long run.

Intra- and interindustry comparisons within the United States underscore the conclusions drawn from the Britain–United States comparison. Haydu demonstrates that where local or industry conditions were closer

to the British situation (i.e., where unions were entrenched prior to the war), the outcomes also paralleled those in Britain (with employers enlisting union support in achieve shop-floor peace as opposed to attempting to sideline unions). Haydu argues that the war opened a window for varying solutions within the U.S. context; however, the window closed after the war when the political balance shifted toward employers and sealed the victory of the open shop.

The strengths of this book are many, and at the top of the list I would place Haydu's skillful use of cross-national and interindustry comparisons that allow him to tease out the interplay between state policy and class relations in shaping labor politics during and after the war. Haydu strikes a balance between analyses that overemphasize the autonomous role of the state on the one hand and those that see state policy as simply a reflection of underlying class relations on the other. Indeed, one of the central achievements of the book is to illuminate the interaction between the two by showing how prewar class relations conditioned what strategies government elites would be able to pursue. For the United States, Haydu offers a nuanced and differentiated account that reveals the important constitutive role played by the state in laying the foundations for employer hegemony after the war. At a theoretical level, Haydu embraces a nondeterministic perspective while at the same time resisting current trends that often overemphasize historical contingency. Haydu shows that while outcomes are not predetermined, the struggles themselves are structured by the political and institutional context in ways that matter for the outcomes (which in turn affect the next round).

Country specialists may quibble with the book's treatment of the "shadow cases," Britain and Germany. Works councils, for example, have somewhat deeper roots in German political economy than Haydu's account—which stresses wartime innovations—suggests (e.g., Hans Teuteberg, *Geschichte der Industriellen Mitbestimmung in Deutschland* [Mohr, 1961]). In addition, more could probably have been said about the coalitional politics (among different segments of business and between segments of business and segments of labor) that influenced outcomes across all these countries. In the case of the United States, Haydu does a very good job of underscoring how employers in different sectors and firms had very different perspectives on the labor question, depending on preexisting labor relations (and this was of course true in the other cases he considers as well). But the argument that links these types of (labor relations–based) differences to political strategies and political coalitions remains somewhat underdeveloped.

These, however, are relatively minor objections, and they should not detract from the considerable accomplishments of the book. This book succeeds on many fronts—as an illuminating new perspective on the origins and evolution of industrial relations institutions and practices in the United States, as a nuanced study of state-society interaction, and as an exemplary application of comparative methodology. As such, *Making American Industry Safe for Democracy* has something important to con-

tribute to a number of fields, including labor studies, political science, sociology, and state theory.

Citizen Politics in Post-Industrial Societies. Edited by Terry Nichols Clark and Michael Rempel. Boulder, Colo.: Westview Press, 1997. Pp. xv+261. \$65.00.

Jane Junn
Rutgers University

The articles in this broad-sweeping volume edited by Terry Nichols Clark and Michael Rempel address the relationship between social and economic forces and political developments in postindustrial society. While the scope is enormous, the focus of most of the articles in this volume is on the attitudes, behaviors, and social and economic circumstances of individual citizens in postindustrial politics. In the introductory chapter, Clark and Rempel detail a general framework for interpreting citizen politics since the 1960s. They argue that during this time period, technology has shifted the politics of postindustrial society, engendering eight core changes, including more influence of the mass media, less influence from social class and religion, declining ideological polarization, rising single-issue politics, greater influence of public discussion, the institutionalization of new social movements, and greater distance between citizens and pressure groups. Further, the editors argue that it is the young and well-educated who are most affected by these shifting politics.

Clark and Rempel divide the succeeding articles into three sections, each of which addresses a more specific question in the broader research agenda in the study of postindustrial politics: Is materialism rising, declining, or both? (pt. 2); What is special about the politics of professionals and other middle class persons? (pt. 3); What mobilizes citizens? (pt. 4). There is a particularly interesting exchange between the authors in part 2. In this section, Ronald Inglehart argues that most industrialized societies show a clear trend toward holding postmaterialist values, while Richard Easterlin and Eileen Crimmins make their case for an increase in private materialist attitudes among young Americans. The chapter by Naoyuki Umemori in this section provides interesting evidence of value change in Japan. Part 3 of the book addresses the emergence and significance of a new middle-class politics. Brint, Cunningham, and Li examine the ideological placement of middle-class professionals in five industrialized societies. In the second article in this part, Brooks and Manza analyze the divergent shifts from conservative to liberal, and vice versa, among groups of workers in the American middle class. The final section of the monograph brings together three articles that address the question of what mobilizes citizens in postindustrial politics. Mayer argues for the increasing relevance of new fiscal populism through his study of political

campaigns since the early 1970s, while the article by Rempel in this section focuses on the ideological cleavages in postindustrial politics. The last chapter in the volume by Butts analyzes survey data from fourteen countries on political activism and feminism.

It is ironic that perhaps the strength of this monograph is also what is its weakness. In casting such a wide net around both the social forces that influence postindustrial politics, as well as the political manifestations of such forces, there is a loss of theoretical focus. At the same time, however, the editors do an admirable job in the introductory chapter weaving together the common theoretical strands of the individual chapters. This collection of articles should be of interest to a wide range of scholars in both the fields of sociology and political science. Students of political behavior, public opinion, social movements, and comparative politics will find the book a useful addition to their libraries.

Modern Manors: Welfare Capitalism since the New Deal. By Sanford M. Jacoby. Princeton, N.J.: Princeton University Press, 1997. Pp. xii+345. \$35.00.

Christopher Howard
College of William and Mary

For many years, scholars have tried to explain why the American welfare state developed later and remained less generous than European welfare states. They have usually stressed what was missing in the United States, such as a strong labor movement and centralized governmental authority. More recent scholarship has pointed out how the comparatively well-developed system of employment-based benefits in the private sector has muted demands for a larger public sector. In particular, several studies have demonstrated how organized labor—in Europe an engine of welfare state expansion—has devoted considerable resources to winning benefits within the framework of collective bargaining in the United States, rather than working wholeheartedly to expand the range and value of public programs.

Modern Manors is an important and timely addition to this new literature. The focus here, however, is how and why companies extended a variety of benefits to workers who were not unionized. Originating in the late 19th century, the practice of "welfare capitalism" ranged from cafeterias, paid vacations, and profit sharing plans to hiring and promotion practices. It included company towns like Pullman, Illinois, and company unions like the Filene Cooperative Association. Jacoby argues that, while we know much about welfare capitalism in the early 20th century, we know little about developments since the 1930s. Previous studies have assumed that the Great Depression obliterated welfare capitalism. Workers realized that they needed greater security against the risks of industrial society than companies could offer and turned instead to the national

government and to unions for help. In contrast, Jacoby demonstrates that welfare capitalism survived the Great Depression, transformed itself, and became a viable option for many firms for the rest of the century.

The heart of the book consists of three rich case histories—Eastman Kodak, Sears Roebuck, and Thompson Products (now part of TRW). The choice of cases is deliberate: the author wants to analyze firms with well-developed and well-documented practices (at times, I was astonished that Jacoby accessed as much sensitive corporate information as he did). Nevertheless, there is sufficient variation to make meaningful generalizations. The three cases illustrate differences between manufacturing and retail firms, between early and late adopters of welfare capitalism, and in the kinds of benefits offered. What united all three companies were desires to forestall unionization, reduce turnover, boost morale, and improve productivity. In addition, at several junctures, leaders of these corporations felt a positive obligation to look after the well-being of their workers; they hoped to create industrial communities of shared interests.

The value of adding a historical dimension to each case becomes evident as each company faces new challenges (e.g., rapid business expansion, new competitors, new labor laws, threats of unionization) and tries to modify relations with its employees accordingly. There are, in short, many data points within each case. Because these challenges were commonplace, firms turned repeatedly to academic researchers and personnel departments to survey workers about their satisfaction with existing programs. Firms' reliance on the emerging behavioral sciences to monitor and sometimes modify workers' demands is one of the more interesting subplots to this history.

My main frustration is that the book provides much greater detail for the 1940s and 1950s than any subsequent period. Given that collective bargaining and unions affect an increasingly smaller fraction of the workforce and that welfare capitalism has become more prominent, it would be interesting to extend the narrative forward. Did the introduction of Medicare and substantial increases in Social Security benefits prompt firms to cut back on benefits? How have welfare capitalists responded to increased global competition? How does the practice of welfare capitalism differ in relatively recent industries like software, management consulting, and temporary services? In Jacoby's defense, he has devoted his energy to the period most often ignored in the literature, and the economics of publishing would probably not have allowed a much longer book. Still, it is hard to know what lessons to draw for the present.

Leaving readers wanting to know more is often the sign of good scholarship, and *Modern Manors* is a very good book. Each case study weaves together developments internal to the company, in the relevant industry, in unionized firms, and in Washington—no mean feat. While the book reflects prodigious research, the author does not divert the reader with pointless facts or minor academic debates. I strongly suspect *Modern Manors* will be at least as valuable to labor historians and business historians as to students of social welfare.

The Hidden Welfare State: Tax Expenditures and Social Policy in the United States. By Christopher Howard. Princeton, N.J.: Princeton University Press, 1997. Pp. xiii+250. \$39.50.

Edwin Amenta
New York University

Most social scientists approach taxes as reluctantly as if they might hasten that other great inevitability. Political scientists and sociologists prefer to address the sunny world of public spending, ceding the gloomy business of taxation to tax specialists and the dismal science of economics. But just as it is impossible to understand life without considering death, it is impossible to understand economic redistribution through social spending without considering taxation. This is especially true for tax "expenditures," commonly known as loopholes or breaks, which reside in the depths of the tax code. That is the fiscal netherworld explored by Christopher Howard, assistant professor of government at the College of William and Mary.

Howard seeks to uncover what he calls America's "hidden welfare state": tax expenditures—departures from the regular tax structure for specific groups or activities—with social welfare purposes. Social welfare purposes in turn are defined by the budgetary categories deemed by scholars to constitute public social welfare expenditures. Some examples are the deduction for mortgage interest for home owners and the earned income-tax credit (EITC), which has been newsworthy in recent years as a means to reform welfare. In his quest to chart the current size and contours of the hidden welfare state, Howard makes extensive use of analytical and theoretical maps for public social spending policy as well as developments in the tax policy literature.

Howard finds that the hidden welfare state is much more sizable than has been generally recognized—constituting about \$400 billion in 1995 or almost half the amount of social welfare outlays. With his discovery, he raises the possibility that American exceptionalism in social spending—commonly understood as relatively low spending for and tardy enactment of social programs—may be a matter of how social welfare is addressed, perhaps through a combination of social and tax expenditures. What is more, U.S. tax expenditures often parallel social expenditures in function, with both giving great attention to old age. The major exception here is in housing policy, for which activity through the front door is greatly surpassed by its back door counterpart.

His mission is also to address the origins and development of the hidden welfare state. He does this mainly by tracing four loopholes across the 20th century. These are cleverly chosen to vary according to whether the tax expenditure is relatively inclusive or means tested and whether it is a cash transfer or an in-kind benefit. Employer pensions and the EITC are both cash transfers; pensions are relatively inclusive, and the EITC is means tested. Thus employer pension loopholes are con-

sidered similar in form to Social Security, and the EITC to the former Aid to Families with Dependent Children (AFDC). The home mortgage interest deduction and targeted jobs tax credit (TJTC) are both in-kind benefits, but the former is inclusive and the latter means tested. Thus the mortgage interest loophole is akin to Medicare and the TJTC to Medicaid.

In his explorations, Howard finds few parallels with the results of social spending studies. Tax expenditures were not typically adopted during the 1930s and 1960s—the decades when American social spending programs made their great leaps forward. For instance, the mortgage interest loophole was carved out in 1913, simultaneous to the introduction of the personal income tax but decades before the Social Security Act. For these reasons, it is not surprising that the standard explanations of the adoption of social spending policy—which rely on Democratic Party dominance of the government, social movement activity, or liberal business influences—do not go far to explain the adoption of the hidden welfare state. Also, the conventional wisdom is that inclusive programs like Social Security are likely to grow, while targeted programs like AFDC face poor long-term prospects. But Howard finds the fates of inclusive tax expenditures is not much different from those of targeted ones. What is more, no beneficiary interest groups have grown up around individual tax expenditures in the way that the American Association of Retired Persons grew up around Social Security.

Howard argues that the difference in determinants between visible and hidden welfare states center on the nature of tax expenditures as a policy tool. He argues that the tax expenditures are linked to taxation policy and are inherently ambiguous. Notably, until 1969, tax expenditures could be adopted without reckoning their probable cost. For these reasons, he argues, tax expenditures are far more easily adopted and less predictable in pattern than typically conflictual social expenditures. Also, Congress has a tighter grip on tax expenditures than on social policy.

Sure as Howard's footing is, following trails blazed by those who study tax policy sometimes leads him astray. Tax experts view loopholes as corruptions of the tax code and thus as conceptually identical to actual expenditures, but this claim should be treated with some skepticism. Counting taxes "lost" through loopholes as if they were expenditures is premised on the counterfactual assumption that without the loopholes the government would have collected the money. As Howard's evidence indicates, however, sometimes the enactment of loopholes constituted part of the process by which permanent tax increases were enacted—Congress's way of sweetening the bitter pill of tax increases, which might not otherwise have passed. Alternatives to loopholes might have included lower rates—which would imply no additional tax collections. In other cases, loopholes may greatly influence behavior, and so in their absence, something other than the collection of taxes might have resulted. Without the tax break on old-age pensions, for instance, employers might have provided higher wages. It makes more sense to discount at various rates

tax expenditures, case by case, rather than merely adding them up at full value. I also wished Howard would have pushed further down the path of comparison to ascertain the degree to which America's tax expenditures constitute a functional equivalent to European spending programs. Finally, tax expenditures differ from social expenditures in more ways than their ambiguity. Social spending programs are mainly redistributive—not the case for many tax expenditures with social welfare purposes.

All the same, Howard's accomplishment is great. He has written what will surely become the official sourcebook to America's hidden welfare state. It is difficult to imagine a better guide through the mysteries of the tax code. Howard cuts through the tangle of jargon and official documents with keen reasoning and lucid style. His employment of theories of and analogies to social policy is always resourceful. *The Hidden Welfare State* shows that scholars concerned with states' efforts toward economic redistribution evade taxation only at their analytical peril.

Labour Market Regimes and Patterns of Flexibility: A Sweden-Canada Comparison. By Axel van den Berg, Bengt Furåker, and Leif Johansson. Lund: Archiv Förlag, 1997. Pp. 264.

Mary Ruggie
Columbia University

This book comes at a particularly opportune moment. Debate continues about the causes and consequences of recalcitrant unemployment rates in certain European countries (Spain, Italy, Germany) compared to the robust levels of employment recently achieved in the United States, the Netherlands, and Britain. Common parlance has it that the rigidities of the labor market in the European welfare states (job protection rules and procedures, powerful and demanding labor unions, and costly social benefits) hinder employers from introducing the proper doses of competition needed to spark economic reform. In contrast, countries that have maintained or introduced greater labor market freedoms are now enjoying the fruits of economic vitality but at the cost of greater social inequality. This position is usually considered to be "neoclassical."

Labour Market Regimes and Patterns of Flexibility addresses the debate about "Eurosclerosis" and its antithesis through an investigation of the relationship between labor market (in)security and workers' flexibility in two countries—Sweden and Canada—that offer an interesting array of differences in labor market relations and policies. At the same time, the geopolitical, industrial, and occupational bases of the two countries are sufficiently parallel to provide adequate controls for comparative analysis. In-depth surveys of employers, managers, workers, and union leaders were conducted in the similarly situated pulp and paper, steel, and telecommunications industries.

Van den Berg, Furåker, and Johansson set their main goal as testing alternative theoretical claims about (in)efficiencies in the allocation of labor. They counterpose the neoclassical understanding to what they call the institutionalist position. It offers both a critique of neoclassical formulations ("the 'free' operation of the labour market does not only lead to inequities but also to inefficiencies" p. 18), as well as a theory of worker motivation: worker security and labor market flexibility "can actually strengthen each other. . . . Increased worker security [can] reduce . . . workers' resistance to change and hence significantly facilitate the process of rapid structural transformation of the economy" (pp. 88–89). The authors also include a variant of the institutionalist position—neocorporatism—which places more emphasis on the role of powerful unions rather than workers, both in resisting change and in negotiating solutions, usually at a tripartite level.

Nowhere in this study are the assumptions of the neoclassical position supported. Job turnover, job tenure, and duration of unemployment are all remarkably similar in the two countries—with qualifications, to be sure. There is some evidence of greater external flexibility in adjusting labor power to demand in Canada, but Canada's higher unemployment suggests less flexibility than in Sweden. Also at the aggregate and behavioral levels, there was no evidence of more resistance to change among Swedish workers, as the neoclassical position holds. In fact, Swedish managers were slightly more sanguine about workers' responses to technical change than their Canadian counterparts.

On balance, the predictions of institutionalist theory fared only marginally better. Canadian workers did not display significantly more hostility toward change. While Canadian workers exhibited, as expected, a greater degree of anxiety about their prospects, this did not translate into any clear pattern of attitudinal differences toward work-related change. However, the slightly more favorable support among Swedish workers for technical change and international competition registered some confirmation of institutionalist tenets.

To round out the authors' surprises, neocorporatist theory predicted quite well differences in how changes were handled at the firm level and in the attitudes of union representatives in the two countries. Swedish union leaders exhibited considerable progressiveness in negotiations with management and in facilitating the process of industrial change.

The authors had the unique opportunity to conduct identical surveys in Sweden at two very different periods—1991 and 1993—and thus test the theories even further. During this interval, Sweden suffered its worst economic crisis since the depression with record unemployment, fairly severe government cutbacks, and deteriorated labor relations. Interestingly, the heightened levels of perceived insecurity that resulted did not affect workers' attitudes toward technical and structural change in industry. The authors conclude that this finding, as well as their general rejection of either the neoclassical or institutionalist perspective, indicates that

labor market flexibility is embedded in the overall sociopolitical context fostered by a set of institutions and arrangements.

While the authors express disappointment at their weak findings, noting that no single theory can account for the patterns they found, perhaps they should not have expected anything else. It is difficult to study the Swedish labor market without drawing conclusions about the powerful role of unions. After reading this book, one is left with the feeling of having learned (or reviewed) a great deal about Swedish labor market relations, less about the economic problems behind Sweden's recently high unemployment rates, and even less about comparable travails in Canada. Ultimately, workers' willingness to change jobs is only one small part of the broader nexus of forces shaping employment; this book gives it a heavier burden of responsibility for structural change in the economy than warranted by the theoretical frameworks.

Economic Ideology and Japanese Industrial Policy: Developmentalism from 1931 to 1965. By Bai Gao. New York: Cambridge University Press, 1997. Pp. xiv+364. \$59.95.

J. Mark Ramseyer
University of Chicago

Born in 1896, Hiromi Arisawa graduated from the University of Tokyo Economics Department in 1922 and joined the faculty as a statistician two years later. He found himself purged as a Marxist in 1938 but rejoined the faculty after the war. He wrote widely, including the 16 books and 31 articles cited in this book. Despite his ideological predilections, after the war he designed Prime Minister Shigeru Yoshida's economic recovery plan and played other prominent advisory roles in the government to boot.

Consider this book the history of the role advisors and bureaucrats played in the Japanese economy, seen through the eyes of Arisawa. As one would expect through Marxist eyes, this is a book that ignores markets. As one would expect through the eyes of a government advisor, it is a book that describes instead how government advisors masterminded economic growth. "You're wonderful," Spencer Tracy apparently once told Katherine Hepburn. Her reply echoes what Arisawa might have said about the work of government advisers: "Of course. Didn't you know?"

Had Gao called his book "Arisawa's vision of Japanese economic history" (something he sometimes seems to do; p. 10), he would have described it accurately enough. He has read widely and deeply in the work of Arisawa and his contemporaries and carefully describes the way they saw their world. As intellectual history, it is a major accomplishment.

Unfortunately, Gao sometimes has bigger ambitions. Rather than intellectual history, he sometimes promotes his study as a history of the role

Japanese bureaucrats and advisors actually played. Given that he focuses almost entirely on what Arisawa and his colleagues wrote, this is more than a bit strange: he uses the writings of 1950s-vintage Marxist consultants not to describe their intellectual milieu but to describe the economy itself. Predictably enough, he describes a world where intellectual masters think great thoughts; where their ex-students in the government dutifully draft statutes and regulations implementing those thoughts; and where businesses meekly do as they are told.

What this ignores, of course, are the dynamics of electoral and economic market competition. Necessarily, the electoral market mediates the transformation of ideas into policy. Intellectuals may talk to bureaucrats, and bureaucrats may like what they hear. But bills do not become law unless politicians pass them, and politicians do not stay in office unless voters reelect them. As a result, whether an idea becomes a statute will usually depend on the effect it will have on the electoral fortunes of the majority party. That the Japanese Diet seemed to implement many of the ideas Arisawa and his colleagues proposed probably suggests that the statutes redistributed wealth to crucial supporters of the ruling party.

Necessarily too, economic markets mediate the transformation of policy into practice. Politicians may pass laws, bureaucrats may make regulations, and both may plead and cajole. But firms will usually do as they are told only if doing so makes money. A government may announce targeted industries, but announcements themselves do not induce firms to alter investment patterns; a government may offer subsidies, but subsidies will not stop firms from trying to take the money and investing it in other projects they thought more profitable all along; a government may mandate cartels, but in industries with many players the cartels will not stop firms from offering hidden price cuts.

None of this is new. Unfortunately for Gao, a massive empirical literature in economics traces the way firms and individuals regularly undercut regulation. To avoid its implications, Gao turns rhetorical: "Methodologically, neoclassical economics derives economic laws through logical and mathematical deduction, disregarding the ethnographic facts about human behavior in any given society and the cross-cultural data about any economic system" (p. 65). By contrast, "the major theoretical propositions of [Arisawa and his compatriots] were derived through induction. They were based on empirical studies of economic reality" (p. 65). Apparently, by declaring modern economics a deductive and historically and culturally contingent ideology ("In my view, Japanese developmentalism [the approach Arisawa took] is simply a competing ideology with Anglo-Saxon economics" [p. 16]), by declaring modern economics theoretical rather than empirical—and by doing this often enough, Gao would wish out of existence those hundreds of empirical studies of regulation. Endlessly, they detailed how ineffective and counterproductive regulation could be. To Gao, Japan is apparently different.

Ambition is a dangerous thing. Not satisfied with an intellectual biography, Gao transforms what would otherwise be a careful and thoughtful

study of the ideas of several 1950s Japanese Marxist academics into a study of the effect of Japanese regulation. Before policy proposals can hit the street, however, they must traverse electoral and economic markets. Both can—and usually do—radically alter the effect those proposals have.

Along the Domestic-Foreign Frontier: Exploring Governance in a Turbulent World. By James N. Rosenau. New York: Cambridge University Press, 1997. Pp. xvii+467. \$59.95 (cloth); \$22.95 (paper).

Tony McGrew
Open University

It might appear a little odd to readers of a sociological journal to confront a review of a work that ostensibly lies within the realm of world politics. But this is a work much more concerned with analyzing the contemporary "human condition" than with the traditional agenda of interstate relations. For in this magisterial sequel to his earlier work, *Turbulence in World Politics* (Princeton University Press, 1990), Rosenau seeks "to confront the insufficiency of our worldviews" (p. 27), which have separated for too long the study of domestic society and the study of global affairs. Indeed the very metaphor of "the Frontier" seeks to capture the growing realm of human affairs which, in an era of intense globalization, transcend the domestic/foreign, internal/external divide. In common with the work of Giddens, Mann, Castells, and others, Rosenau delivers an account of global social change. This is premised on the proposition that to understand the contemporary human condition, "the analyst has little choice but to probe both the internal and external dynamics of societal life and the intricate connections between them" (p. xiii). A terrifically ambitious project, it is executed nevertheless with enormous conviction and imagination.

The original essays that form the foundation of this book have been revised significantly and combined with additional original material to produce a cogent and coherent analysis of "life on the Frontier." But this adds up to more than the author's modest claims of a "set of explorations rather than a definitive work" (p. xiii). For what binds the analysis together is a distinctive account of global social change that amounts to a new way of thinking about social and political life in a globalizing world. In this respect, the work is both definitive and also hugely relevant to the study of social change.

Central to his account is the conviction that societies and world politics are undergoing a profound transformation. Both are in the grip of contradictory forces, and at the same time, they are having to adapt to the consequences of growing global interdependence and the growth of what Rosenau refers to as "sub-groupism," the desire for autonomy as peoples seek greater control over their own individual or collective destiny. As

Rosenau comments, "To analyze globalization and localization processes, in short, is to observe the world in motion" (p. 85). That dynamic motion he argues is best understood as "framnegration," an awkward syllogism but one that seeks to capture the contradictory imperatives of global integration and societal disintegration. "Framnegration" not only defines the contemporary human condition but also underlies the growing intersection of domestic and foreign affairs, the expanding "Frontier" that constitutes "a widening field of action . . . the space in which world affairs unfold" (p. 5). But not just world affairs, for "one is hard put to conceive of any issue or aspect of the social and cultural life of communities at all levels not being modified by the dynamics of framnegration" (p. 108).

Of primary concern to Rosenau is the question of governance: how "framnegration" is transforming governance within and between societies. In this more turbulent world, so the argument goes, the state is increasingly becoming one amongst many sites of authoritative decision making and order. Authority and governance is being relocated upwards, sideways, and downwards within supranational, transnational, local public and private "spheres of authority" respectively. These spheres not only transcend national territorial jurisdictions, creating new configurations of social and political space, but also involve a multiplicity of interested agents, from governments to Greenpeace. Such developments, as Rosenau argues, prefigure profound challenges to orthodox notions of democratic accountability, legitimacy, representation, state sovereignty, territoriality, national identity, and citizenship. In other words, they imply the "reinvention" of politics itself.

Although the author presents an innovative and compelling analysis of global social change (alongside its implications for patterns of global governance), it is nevertheless subject to qualification. In particular, there is a curious absence insofar as there is little systematic discussion of global markets and transnational capital despite their centrality to the argument. Thus, while there is an assertion that states and peoples are losing out to global markets, this issue is never convincingly addressed. Moreover, despite the appeal for a new ontology of global affairs, the analysis retains significant traces of a liberal and a pluralist view of politics and governance. Such implicit starting points condition judgments about the political consequences of "framnegration" and the prospects for governance along the "Frontier" insofar as the analysis tends to overlook the systematic examination of inequality and hierarchy inscribed in processes of globalization and localization. These silences are curious in a work that is conscious that "the dangers of oversimplification far outweigh those of excessive elaboration" (p. xiii).

Throughout his prolific academic career, Rosenau has always pioneered the intellectual "Frontier" confronting the skeptics and the orthodoxies of the day. This work is no exception. It is essential reading for anyone interested in the forces that are transforming modern societies at the end of the 20th century.

The Challenge of Interracial Unionism: Alabama Coal Miners, 1878-1921. By Daniel Letwin. Chapel Hill: University of North Carolina Press, 1998. Pp. xii+289. \$49.95 (cloth); \$19.95 (paper).

John Lie

University of Illinois, Urbana-Champaign

Sociologists are wont to bypass a narrative history of Alabama mine workers. They should not. Given the centrality of race and class in contemporary sociological debates, Daniel Letwin presents a rich and revealing history of the intersection of the race and class lines that resist the received, hackneyed narrative of Southern, working-class racism.

Bereft of water but blessed with minerals, boosters sought to transform Birmingham, Alabama, as the "Pittsburgh of the South" in the late 19th century. The rise of the mining industry brought white and black convict labor and part-time farmers to the Birmingham area. Letwin traces the diverse origins, paths, and categories of workers, as well as their sources of solidarity and schism. Although the author ably limns the fluctuating fortunes of the mining industry and the political economy of Alabama from 1878 to 1921, his central interest is in the trajectory of miners' unions—from the Greenback-Labor Party to the Knights of Labor to the United Mine Workers—and their efforts to forge interracial solidarity.

In recounting cycles of promise and failure for unions, Letwin underscores several themes of great interest. First, he emphasizes the persistence of interracial solidarity in the culture of white supremacy; labor unity overstepped the color line time and again. Black and white workers resisted, among others, the use of convict labor, the influx of contract labor, and the introduction of strikebreakers. Not only racial interest but shared experience and ideology bound white and black workers together. Even at the time of heightened racial hostility and the retrenchment of populist and progressive politics in the 1900s, interracial solidarity remained robust.

Second, Letwin delineates the limits of interracial solidarity. White miners by and large presumed white supremacy. Not surprisingly, mine operators used the "race card" repeatedly to challenge racial solidarity; the threat of "social equality" strengthened the color line at the expense of the class line. More interestingly, he argues that interracial solidarity succeeded because there were not women miners. The threat to the "sanctity of white womanhood"—that is, black men interacting with white women—undermined racial solidarity. In other words, interracial solidarity could resist the race card but not the gender card.

Third, Letwin suggests an interesting model of labor activism and defeat. The upsurge in union activities inevitably generates mine operators' resistance. The spiral of strikes and strikebreaking engenders racial rift (mine operators' use of the race card) that undermines not only labor solidarity but public support. Ultimately, labor's recourse to violence saps support for unions.

The Challenge of Interracial Unionism offers, of course, much more. It recovers little-known movements, such as the Greenbackers, and many moments of interracial solidarity, all too easily overshadowed by the culture of Jim Crow. It has, needless to say, some shortcomings. Most importantly, it does not quite succeed in explaining the rise and fall of interracial unionism. I think he would need to consider more seriously both the macrostructural context and the concrete social networks. But Letwin's is an interesting and inspiring book that forces us to rethink the fundamental issues of race and class in the United States, past and present.

Paths to Success: Beating the Odds in American Society. By Charles C. Harrington and Susan K. Boardman. Cambridge, Mass.: Harvard University Press, 1997. Pp. x+238. \$35.00.

Jennifer L. Hochschild
Princeton University

Most anthropologists, as the authors of *Paths to Success* point out, do not "study up"; most psychologists analyze only pathological forms of deviance; most sociologists are much better at explaining why members of a given population are likely to fail than to determine how a few might succeed. Yet Americans are bred on the Horatio Alger myth, and millions of people have for centuries emigrated to the United States in order to pursue the American dream of success. This book begins to close the gap between popular ideology and scholarly attention by closely examining the lives of unusually successful black and white men and women. Although it attends to issues of race and gender, its key distinction is between "pathmakers" and "controls." The former "are those with successful careers in high-status occupations whose families of origin were poor . . . , low in occupational status, and whose parents were not high school graduates" (p. 26). The latter are equally successful but come from higher-status backgrounds. Starting in 1982, the authors and their associates investigated 60 pathmakers and 40 controls, using a three- to five-hour structured life history interview. The initial subjects were attained "through the social networks of the senior researchers," and these subjects suggested others. The goal of the interviews was to determine how pathmakers and controls resembled and differed from one another in their route to success.

Harrington and Boardman focus mostly on standard sociological variables in tracing that route, seeking to determine the effects of families of origin and generation, schooling, mentors, peers, and first jobs on their respondents' life course. They also consider classic psychological variables such as need for achievement and power, locus of control, and maturity of defenses. They rely on several key concepts, such as spiralism (a small success generates another, slightly larger), redundancy of resources (whites, men, and the well-off have more to fall back on if one path to

success is blocked), and the idea of pathmaking itself. Their results are interesting though hardly surprising: controls enjoy more redundancy of resources than do pathmakers; families of origin sometimes help and sometimes hinder the effort to succeed; a teacher or other mentor can have a crucial impact on the life of a child with few other resources; successful middle-aged women are less likely to be married or to have children than men of comparable status. The real value of the book lies in the extended quotations from eight of the subjects who were chosen to exemplify the rest. They speak eloquently, bravely, insightfully, and sometimes heartbreakingly of the barriers they have overcome as well as those that still block them and of how they have attained at least some portion of the American dream.

I wanted very much to like this book, but I was unable to muster much enthusiasm except for the subjects themselves. I see several flaws. First is the fact that the book is organized around the standard statistical no-no of selecting on the dependent variable. (The authors briefly justify that choice, but not convincingly in my opinion). In addition, the method of selecting respondents seems unnecessarily adventitious. Second is the fact that, however they attained their respondents, they did not make as much as they could have of the interview material. As Harrington and Boardman point out several times, the book focuses on description rather than explanation or interpretation on the grounds that we know so little about successful people that they need first to lay the groundwork for hypothesis testing. I have no direct quarrel with that argument; my disappointment lies instead in the fact that they did not do as much with the wonderfully rich set of interviews as one could have. For example, interviewing started 15 years before the book was published. What happened to the subjects in the interim? Did women or African-Americans attaining success in the mid-1990s have different experiences from those attaining success in the rather different political context of the 1980s? I am a political scientist, so I cannot resist asking about the political ideologies of the pathmakers. Were they grateful to America for giving them the chance to succeed or bitter about how much more difficult it was for them than for others? Do the white men (especially the controls) feel defensive about their relative advantages? How did the African-Americans feel about the growth of the black urban "underclass" of nonachievers? And so on. Overall, the interview topics and analyses feel rather unimaginative and cut-and-dried; the reader has many fewer "ah, ha!" experiences than one would hope for with this terrifically interesting sample and extensive array of qualitative material.

I thus deem this a worthy endeavor rather than a complete success. If it stimulates more of the same kind of research, that will be terrific, but by itself *Paths to Success* leaves at least this reviewer a bit flat.

Against the Odds? Social Class and Social Justice in Industrial Societies.
By Gordon Marshall, Adam Swift, and Stephen Roberts. New York: Oxford University Press, 1997. Pp. xiii+265. \$75.00 (cloth); \$29.95 (paper).

Charles E. Hurst
College of Wooster

Against the Odds consists of a detailed empirical and theoretical examination of the relationships among one's classes (occupational status) of origin and of destination, education, meritocracy, and social justice. Social justice is defined as equality of access. In this sense, the book is part of the larger body of research that concerns the legitimacy and openness of class structures and the processes involved in social mobility.

The authors provide clear definitions of their central concepts and construct their arguments in detailed, systematic fashion. They recognize, for example, the flaws in concluding that educational attainment necessarily reflects merit or that a close tie between such attainment and class of destination means that the class structure reflects a meritocracy. The greatest strength of this volume is the care taken in discussing (a) the substantive issues involved in the meritocratic triad (i.e., the linkage between classes of origin and destination and the mediating role of educational attainment) (b) the many meanings of "meritocracy" and the myriad of philosophical and theoretical problems involved in deciding whether a distribution process is just, and (c) the relevance of their British data on class mobility for answering questions about social justice. I found their intricate discussions of the problems in determining the presence of meritocracy especially rewarding. Happily, several appendixes are provided for those unfamiliar with the broader research context and the statistical techniques used in the study.

Being concerned with access, the authors begin with an examination of the industrialism thesis using data from 11 countries and conclude that there is little evidence to support the belief that higher mobility rates or greater equality of opportunity are found in more industrialized countries, including Britain, or that there is a convergence in mobility among such countries. Nor does there appear to have been significant change in inequality of opportunity throughout much of this century. The reader should be cautioned that the most privileged class categories are not included in the analysis and, for comparative purposes, some class categories are collapsed into more generic groupings.

The persistence of this inequality raises the possibility of social injustice, which is the subject of the remainder of the book and its greatest strength. In analyzing the *possibility* of meritocratic processes at work, the authors test the three legs of the meritocratic triad. The first general conclusion is that class of origin continues to affect class of destination directly for both men and women in all industrial countries, including Britain, regardless of educational attainment. On the second relationship, between class of origin and educational attainment, the gap between

classes in reaching the higher levels of education in Britain has increased over time, which would appear to weaken the chances for those from lower classes to reach higher occupational statuses. Some of the trends presented are curvilinear but are interpreted largely in linear terms. A corresponding explanation of fluctuations in broader economic conditions would have helped clarify what appear to be nonlinear trends. On the third relationship, between educational attainment and class of destination, it is found that, similar to other studies, while there are some variations at specific educational levels, the general relationship between educational attainment and class of destination in Britain has not changed much during the 20th century, suggesting no overall increase in merit selection or social fluidity. There also continues to be a significant, but weakened, direct tie between classes of origin and destination.

The authors then walk into the "minefield" (p. 134) of relating their findings on social mobility to the issues of meritocracy and social justice. They relate the difficulties of associating educational attainment with merit, of identifying merit associated with different classes that are not reflected in education, and of choosing between varying conceptions of merit. Because of the complexities surrounding the notions of merit and desert (e.g., issues of responsibility, preferences, market forces, motivation, luck, etc.), a definitive conclusion about meritocracy cannot be reached in the book. On the subject on the factors over which individuals do or do not have control, the theoretical and philosophical analysis could have benefited by references to ethnographic studies of class reproduction. The authors conclude that, while it may be plausible to conclude that the British get jobs on the basis of their competencies, it is less secure to argue that people of varying competencies should be differentially rewarded because of the many "nondeserving" sources of competencies. They raise a number of issues, many of which can be empirically examined and which must be addressed before firm conclusions about the justice of reward distributions can be made. In this sense, the authors lay the groundwork for further research on issues of meritocracy and social justice. While too complicated for undergraduates, their study would be especially useful for graduate students examining stratification processes and status-attainment researchers attempting to enrich their models.

Capital Culture: Gender at Work in the City. By Linda McDowell. Malden, Mass.: Blackwell Publishers, 1997. Pp. xvi+240. \$24.95 (paper).

Amy S. Wharton
Washington State University

Globalization, as an economic, social, and cultural phenomenon, has received considerable attention from scholars. Interest in the social and cultural dimensions of this process—especially as it affects the workplace—is reflected in research on the expanding service economy and the chang-

ing demographics of workers. Studies of banking and finance, on the other hand, have provided insight into the economic aspects of global capitalism. Missing from these literatures, however, is attention to financial work, financial workplaces, and financial workers "who ma[k]e their living by moving invisible sums around the globe or by advising companies to invest, divest and take over other firms whose products and workers would never be seen" (p. 159). As a critical component of global capitalism and the service sector, finance represents an important vantage point from which to explore work, culture, and society at the end of the 20th century.

In her book, *Capital Culture: Gender at Work in the City*, Linda McDowell contributes to an emerging literature on these issues. As a geographer, McDowell may be especially well suited to explore the relations between the spatial and the social that are at the heart of a global economy. By drawing on what she refers to as her "sociological and geographical imaginations," McDowell offers readers a theoretically-driven, empirical analysis of gender, power, and space in the globalized, financial workplace (p. 12). Focusing on three investment banks in the city of London, McDowell is particularly interested in understanding the processes that sustain gender segregation in these settings.

The book begins with a provocative theoretical discussion of the multiple literatures that inform the author's work. I highly recommend this section to researchers grappling with ways to conceptualize the changing workplace. McDowell covers some ground that will be familiar to students of gender and service work, but she embeds these observations within a larger discussion of what she calls the "cultural turn" in recent studies of work and organizations (p. 12). This turn is one "in which the perception of work and workplaces as active forces in the social construction of workers as embodied beings has become a prominent emphasis" (p. 12). McDowell combines this sociological notion, which has been used to understand how women and men are constituted as gendered beings inside the workplace, with insights from geography and architecture. These latter fields are drawn on to illustrate how "the location and the physical construction of the workplace . . . also affects, as well as reflects, the social construction of work and workers and the relations of power, control, and dominance that structure relations between them" (p. 12).

This theoretical discussion is followed by empirical analyses designed to reveal the social and spatial contexts that foster gendered work practices in investment banking. In a fascinating chapter on the physical and social organization of finance in the city of London, McDowell describes the recent "Americanization" of these financial establishments. This transformation is marked by a gradual deemphasis on family background and inherited status as factors necessary for success and a growing importance on technical and professional expertise. This shift in the culture of finance and the meanings attached to work is also reflected in the city's changing built environment. Most notably, McDowell notes: "Instead of money-making being serious work undertaken by men in spe-

cialized and exclusive spaces, service sector work was redefined as fun" (p. 60). For middle-class members of the financial sector, the spatial segregation between sites of work and play and between public and private life is being reduced.

The middle portions of McDowell's book focus on gendered work patterns and career paths in the three banks. Women workers in these settings were much more likely to work in clerical positions than professional or managerial jobs, and all three banks were characterized by a high degree of gender segregation in patterns of interaction. Drawing on interview and survey data collected from workers in each bank, McDowell explores women's and men's career paths and the ways in which the culture of banking reproduces social class and gender distinctions. While these chapters provide an interesting glimpse into the gendered world of investment banking in the city of London, I found them less compelling than the book's later chapters, where McDowell explores the construction of gendered selves and bodies at work and returns to the theoretical issues raised early on.

A central claim in these later chapters is McDowell's contention that service sector work, including finance, is "leading to the 'feminization' of all workers in the sense that bodies and personal appearance have become an integral element of workplace success" (p. 139). This notion, together with her earlier observations on the disintegrating boundary between work and leisure, frame McDowell's analysis of how women and men in finance "do gender." She argues that older versions of "patriarchal masculinity" may be disappearing from the financial workplace and gender identities and interactions at work are "within bounds, fluid and negotiable" (p. 208). Nevertheless, McDowell concludes that the feminized, service-sector workplace—in finance, at least—continues to enforce gender distinctions in ways that reproduce gender inequality.

This book deserves a wide audience: students of the service sector should find McDowell's theoretical and conceptual insights about this topic useful, students of gender and work will encounter a carefully drawn case study of how gender distinctions are constructed and reproduced on the job, finally, those interested in cultivating links between their sociological and geographical imaginations will find that *Capital Culture* can help them achieve this goal.

Liberty, Equality, and Justice: Civil Rights, Women's Rights, and the Regulation of Business, 1865-1932. By Ross Evans Paulson. Durham, N.C.: Duke University Press, 1997. Pp. vii+361. \$59.95 (cloth); \$18.95 (paper).

Eileen Boris
Howard University

This interpretative synthesis describes key reform movements of the late 19th and early 20th centuries by situating discourse in its economic, political, and social context. By tracing the shifting meanings and "relative ranking" (p. 4) of liberty, equality, and justice—defined as this nation's core values—historian Ross Evans Paulson presents an exercise in applied sociology that is more successful as a historical survey of struggles for civil rights, women's rights, and the regulation of business. He adapts categories from Robert K. Merton to characterize responses to a growing disjuncture between core values and social practices. Resigners accepted the status quo; ritualists withdrew into their own routines for personal renewal; reformers strove to close the gap between values and social behavior; and radicals sought new values and alternative institutions. But this is no mechanical application of theory—in the guise of "situational analysis, historical semantics, and methodological caution" (p. 14)—to what Paulson recognizes as the contingency, flux, and difference of the past. With social language as the subject, core values become the moving force of history. Reformer failure derives "primarily from the fact that so many Americans ranked liberty (for themselves) higher than equality (with others) and justice (for all)" (pp. 1-2).

"Part I: Old Languages and New Realities" explicates what happened to inherited understandings of liberty, equality, and justice after the Civil War, Progressivism and its aftermath in the consumer-oriented 1920s forms the subject of "Part II: New Languages and Old Realities." This section illustrates the social languages of antimonopoly, social cohesion, and efficiency identified by historian Daniel T. Rodgers in 1982 but here linked both to organizations and politics and to the social ends of reformers, radicals, and ritualists. If theoretical models sometimes become lost in a wealth of details, at other times Paulson deploys them brilliantly, as when comparing Garveyism with the Ku Klux Klan. Though both appear as ritualistic responses, Garveyism gave core values "radically new meanings in terms of extranational racial citizenship, black nationalism, racial pride, and group-based self-help projects" (p. 242). In contrast, the Klan undermined the very values it evoked.

The free labor ideology of the Republican Party had emphasized liberty as independence. Economic individualism embraced notions of liberty as "the *freedom to* pursue your own interest and the *freedom from* external restraint" (p. 19). Equality would generate liberty by bringing all under the auspices of the law. Justice appeared "as a matter of strict constitutional limits, specified procedures, and infrequent appeals to higher

courts" (p. 21). In Paulson's scheme, congressional Reconstruction was reformist, even though it appeared radical because key Republicans emphasized equality over the other core values. But by leaving land in the hands of the exconfederates, they ended up restricting the liberty of the freed people.

Through use of the 14th amendment, the modern corporation certainly embraced the language of liberty. However, its behaviors "seemed to diminish the reality of both liberty and equality, both competition and cooperation, both individual mobility and community autonomy," legal privileges as well as "equal protection of the law" (p. 236). Neither the languages of antimonopoly nor efficiency could mitigate such costs. Paulson considers the alternative visions of organized workers and socialists, but his framework underplays the reality of economic power.

Though recognizing that social historians have uncovered "a tale that is more negative than positive" (p. 39), Paulson rarely compares public utterances with more private ones or those of leaders with those of ordinary people. Later chapters, for example, introduce the diversity of black political thought, discussing such figures as Thomas Fortune, Mary Church Terrell, and Ida B. Wells-Barnett as well as Booker T. Washington, W. E. B. Dubois, A. Philip Randolph, and Garvey. But Paulson never captures the rights talk of the black masses. So the story of civil rights becomes one of judicial and legislative decisions. African-Americans gained freedom from servitude without freedom to act as they saw fit. Other minorities faced an even bleaker situation, with Chinese excluded from citizenship and Native Americans promised equality only when they moved from communalism to individualism.

Paulson particularly admires those who emphasized interdependence—Native Americans, African-Americans, working class ethnic communities, middle-class white maternalists, and male socialists. Thoroughly conversant with recent work in women's history, he builds upon his earlier studies of women's rights and prohibition to explicate the tension between dependence, independence, and interdependence in women's lives. His deft discussion of manhood among social gospel advocates illuminates the ways that gender consciousness pervaded reform; his presentation of the equality vs. difference debate among women advances our understanding of how individualism captured the mantle of feminism, trumping the language of community and social justice. Similarly, the modern concept of civil liberties illustrates how individual liberties became "*primary*" (p. 185). Specialists know this history, but Paulson offers fresh insights from probing contested key words to capture cultural conflict within individuals as well as institutions.

Creating Socialist Women in Japan: Gender, Labour and Activism, 1900–1937. By Vera Mackie. Cambridge: Cambridge University Press, 1997. Pp. ix+252. \$54.95.

Sharon Sievers
California State University, Long Beach

This is a very ambitious effort to explore the “creation” of socialist women in modern Japan. Mackie’s strategy is to analyze, in a historical context, the expressions of women in a number of different genres as they moved from theorizing a class position to helping build institutions that would make a socialist movement viable; Japanese women, Mackie tells us, had to transform their own subjectivity and ultimately imagine themselves as workers, comrades, and activists rather than Imperial subjects, wives, and mothers. How difficult a task this was in modern Japan is clear from Mackie’s recounting and analysis of the historical context in which Japanese women functioned.

Each of the early chapters address “speaking positions” open to women: (Imperial) subjects, wives, mothers—all in service to the imperial state but all capable of moving to new subjectivities as workers and activists. Each of these chapters follows a rough chronological framework, and each addresses three sets of questions: (1) What were the theoretical concerns of socialist women, and how did they understand the relationship between socialism and feminism? (2) How were these ideas translated into action, and what forms of organization and action were developed? and (3) What imaginative resources were available for imagining the transformation of society, what rhetorical strategies and metaphorical tropes were employed, and how did these translate into new forms of gendered subjectivity (p. 21)?

The subtext of many of these questions is the troubled and complex relationship between feminism and socialism, both born of 19th-century liberal tendencies, the problems women had as they tried to bring feminist issues to the Japanese socialist movement, and the frustration many felt in their efforts to raise class consciousness among many Japanese feminists.

It is clear why Mackie would use Fukuda Hideko’s *Half My Life* as an entry point to these discussions. This autobiography is unusual in its capacity to explain how Fukuda developed a political identity and how she struggled to be both woman and activist in late 19th-century Japan. The problem is that few such clear cut literary records exist; though women like Miyamoto Yuriko and Sata Ineko, both active much later in left-wing politics, wrote interestingly and informatively about some of these same issues, they too provide what are rare glimpses into the lives of activist women. In the face of this reality, Mackie is forced to rely on a number of different sources, some written by women, some not. It is unfortunate but true, for example, that the early socialists in the Commoners Society wrote a great deal, given their small numbers. But women

in the Commoners Society wrote relatively little compared to men in the group.

What is clear from the record is that Japanese women, after a hopeful beginning in liberal political movements, found themselves increasingly absent from the political arena. The government, through constitutional law, education codes, and above all, a new Civil Code, had made Japanese women virtually invisible. Under regulations formulated in 1890, women were completely excluded from politics; they could not vote, organize, or associate with political groups. So, from Mackie's point of view, the state's effort to include women on its own terms, first as "good wives" then as "wise mothers" in a patriarchal family-state structure, simply meant that the "official construction of women . . . positioned [them] as passive recipients of state protection" and "more or less passive supporters of a military state" (p. 71).

But clearly the story does not end there. Women who, as members of the Commoners Society at the turn of the century, might have agreed with the essentialist descriptions of their roles enunciated by male comrades and who wanted to create a family of comrades had, by 1920, developed what Mackie describes as the ability to imagine themselves as workers and activists.

"Gaining in assertiveness was one aspect of the transformation of the identity of 'factory girls' into an identity which gave equal weighting to the fact that they were workers, members of the working class, and potential unionists" (p. 122). Still, what women needed in this circumstance was an image of themselves that recognized class differences among women and emphasized their strength in solidarity. Mackie quotes Tajima Hide: "It is the factory girls who are doing the most important work. If all of the 730,000 factory girls were now to stop working, we would find out how much power we have. Until now we have been unaware of our power . . . Let us show our strength, to ourselves, and to the world. Let us be proud of our strength" (p. 123). From 1929 to 1931, in the midst of worldwide economic dislocation, it was women textile workers who still constituted a majority of the Japanese workforce, whose militancy made strikes like the great Toyo Muslin Strike of 1930 possible. Finally, Mackie tells us, their numbers and their militancy began to open the eyes of male unionists to their potential as workers and comrades in the movement.

Does Mackie tell us everything we would like to know about the creation of socialist women in Japan during the turbulent period from 1900 through 1937? Probably not. But she has made the most of sources available to her and has presented us with a very clear and thought-provoking look at the transformation of working women in Japan in this period.

Real Knockouts: The Physical Feminism of Women's Self-Defense. By Martha McCaughey. New York: New York University Press, 1997. Pp. xvi+270.

Bonnie H. Erickson
University of Toronto

Can women, and should women, forcefully defend themselves when attacked by men? McCaughey argues that the answer "no" is central to gender ideology. Women are supposed to be too weak and too nice to fight back, a view with dreadful costs for women: they learn to be afraid, vulnerable, and submissive. Not only popular thinking but some strands of feminism assume that men are stronger and more aggressive or argue that women ought not to participate in the male vice of violence even for self-defense. Women's self-defense courses can powerfully subvert such gender ideology by physical, compelling experience; women find that they really can fight and even have fun doing it.

McCaughey learned about women's self-defense through 120 hours of participant observation in several kinds of classes, interviews with some teachers and students, and inspection of some self-defense literature. From this material, she has crafted an engaging book, lively in language and full of feeling, which opens up some pivotal issues for feminists and self-defense advocates alike. Feminists need to more fully recognize the ways that gender scripts are written into and through bodies and the value of bodily experiences and practices that can erase such scripts in favor of new empowering ones. Some self-defense practitioners are already gender subversives, but many could be more feminist in their ideas and teaching practices.

McCaughey makes arguments that often ring true to me as a longtime martial artist, putting these themes into an appealing book-length form that will make them more widely known and, I hope, more systematically analyzed and investigated. McCaughey's own empirical material is vivid but modest in scope and not at all designed to test fundamental questions: whether training has the effects described, how it does if it does, and who gets to benefit from training.

Who benefits? McCaughey (p. 216) could not find (U.S.) national data on self-defense course participation, but the 1988 General Social Survey of Canada asked, "During 1987, did you do any of the following things to protect yourself or your property from crime . . . taken a self-defense course?" Participation was low (3%), much rarer than installing home security (23%) or changing routine activities (25%). Consistent with women's learned reluctance to fight, self-defense training was more common for men than women (3% vs. 2%); consistent with the cost of training, it was more common for the rich than the poor; like many forms of physical exercise, it was more common for younger people; and it was most common among people who had already been victimized, especially if vio-

lently victimized (Vincent F. Sacco, *Patterns of Criminal Victimization in Canada* [Statistics Canada, 1990]). Self-defense training is too rare, too unequally available, and often too late.

Some of the less privileged learn the powerful bodily lessons of self-defense the hard way: learning by doing. Some lower-status jobs, from night shift at a convenience store to security guards and investigators, include routine exposure to risk. My interviews with women security workers in Toronto show that some of them have defended themselves successfully under tough conditions and developed a new sense of power, discovered the thrill of action, and in short, found the same benefits McCaughey describes for self-defense classes. But some women, untrained for the risks, have been badly injured. Few security employers give self-defense training, and only four out of 10 guards or investigators have training from any source.

Does it matter how women study self-defense? For example, is it better for women to train with both men and women (so women can learn to handle men and prove that ability in practice) or to train in women's programs (which can directly address issues such as past assault experience or reluctance to risk hurting anyone)? McCaughey clearly has her favorite methods but cannot assess them with her own materials. There is some other published research that makes points very similar to hers (such as S. R. Guthrie's 1995 report on a feminist karate club [*Women and Therapy* 16:107-19]) and a series of preliminary impact assessments in back volumes of *Perceptual and Motor Skills*, but the research is all in places with very low visibility for most sociologists and is usually fairly limited in scope and design. There is a huge amount of work needed on various ways of teaching and training and their impacts on an array of possible outcomes (feeling strong, becoming strong, gaining self-confidence, gaining insight into gender ideology, becoming able to defend oneself, and so on) as well as how long-lasting such impacts may be.

This work will be important for our understanding of gender, bodies, and culture, as well as important for women's safety and self-actualization. McCaughey's passionate and persuasive book does a great service in opening up these issues to a wider audience.

Media and Political Conflict: News from the Middle East. By Gadi Wolfsfeld. New York: Cambridge University Press, 1997. Pp. xiv+255. \$59.95 (cloth); \$18.95 (paper).

Darnell M. Hunt
University of Southern California

The role of news media in political conflict continues to be a hotly contested issue. From the Right, scholars often charge news media with harboring a liberal bias and with undermining traditional values and authority. Scholars from the Left, in contrast, routinely identify the media as

transmitters of conservative ideology and as integral players in the maintenance of the status quo.

In *Media and Political Conflict: News from the Middle East*, Wolfsfeld attempts to reconcile the two positions. He selects three case studies to illustrate his proposed model for explaining the various roles media play in political conflict. The first case study reviews the struggle over the Israeli-Palestinian Declaration of Principles signed in September of 1993 (the Oslo Accords); the second focuses on the low-grade Palestinian uprising against Israeli troops that began in December of 1987 (the *intifada*); the final examines the Persian Gulf conflict that ensued after Iraq's invasion of Kuwait in August of 1990. Underlying this choice of case studies was Wolfsfeld's goal of challenging "static" conceptions of media that present them as *either* "faithful servants" of the power structure *or* as "advocates of the underdog" (p. 69). Indeed, Wolfsfeld convincingly uses personal interviews with journalists and political leaders from each of the cases, as well as analyses of the dominant frames shaping important news texts about the cases, to locate the media's shifting role on what he describes as a "continuum of influence" (p. 69). While the media's promotion of Allied frames in the Persian Gulf conflict defined the "faithful servant" end of the continuum, for example, the media's acceptance of Palestinian frames of "injustice and defiance" in the *intifada* defined the opposite extreme. The media's role in the debate over the Oslo accords—that of "semi-honest brokers"—defined a middle position on Wolfsfeld's continuum (p. 69).

Wolfsfeld premises his analysis on the assumption that the news media have become "the central arena for political conflict" (p. 2), one where political antagonists (i.e., authorities and challengers) compete for public acceptance and where journalists serve to turn "even the most monotonous of contests into exciting drama" (p. 1). There are no foregone conclusions in this struggle. Just as Wolfsfeld rejects models that simplistically predict hegemonic outcomes for media influence, he also seems to dismiss models that attribute counterhegemonic outcomes directly to the politics of individual newsmakers or news organizations. For Wolfsfeld, the media's position on the influence continuum in any given case is more appropriately understood as the joint product of two important structural (i.e., situational) factors: the power that political antagonists wield over the news media *and* the power that news media are able to wield over political antagonists. Indeed, he proposes a rather simple formula for calculating these power differentials: "power" for political antagonists is defined as the ratio of their value to the media over their dependence on the media; conversely, "power" for individual media outlets is defined as the ratio of their value to the political antagonists over their dependence on the antagonists.

When a political antagonist wields a disproportionate amount of power over the news media (e.g., the Allied forces monopoly of battlefield information during the Gulf War), for example, news media are reduced to uncritically accepting the antagonist's framing of events. However, when

news media have a great deal of power relative to political antagonists (e.g., the news media's easy access to images of the *intifada*; Israeli dependence on the media in its campaign to promote a positive image to the United States and other powerful allies), news media are better positioned to influence the tactics, strategies, and political standing of the antagonists. As a consequence of this power advantage, Wolfsfeld concludes, news media are also better positioned (and more likely) to circulate news frames that challenge those promoted by authorities. The cultural dimension in which these framing choices are made is defined by three basic questions (p. 49): "How did we cover this conflict in the past? What is the most newsworthy part of the conflict? Who are the bad guys?"

While I applaud Wolfsfeld's specification of important situational determinants of power in the case studies, his apparent decoupling of power and ideology seems somewhat curious. Nowhere in Wolfsfeld's exchange-theory-based power models does the issue of ideology arise. Instead, ideology is relegated to the "cultural dimension" (p. 31), a realm that merely "overlap[s]" (p. 218) the "structural dimension" (p. 13) that he claims *directly* determines the relative power of media and political antagonists. But Wolfsfeld's own case studies (as well as an impressive body of scholarship) inadvertently reveal that these dimensions more than *just* "overlap." They are intricately interconnected. For example, Wolfsfeld's observation that U.S. authorities had more power than news media in the Gulf Conflict (i.e., the news media were dependent upon the military for updates from the battlefield) presupposes that "news" from the battlefield was *the* important "news." What about the sizable protest movement that emerged in the United States surrounding the conflict? Could it be that the media's "dependency" on military authorities had less to do with the military's control of information from the battlefield than with the media's ideological proclivity (i.e., journalistic routines and news values) to define military conflict as more "newsworthy" than civilian protest? In the end, the concept of hegemony—which Wolfsfeld mentions several times in passing but often derisively (e.g., "Hegemony indeed!" p. 178)—fleshes out the very real links between power and ideology. Despite its failure to more fully explore these links, *Media and Political Conflict* is loaded with insights that will undoubtedly spark further research.

Music and Cultural Theory. By John Shepherd and Peter Wicke. Malden, Mass.: Polity Press, 1997. Pp. ix+230. \$32.95.

Robert W. Witkin
University of Exeter

When my daughter was three years old, I watched her eyes fill with tears whenever the Christmas carol "Away in a Manger" was hummed. Even without the words the tears came. Was this a learned response to the

culturally mediated communication of affect—to conventions that attach affective import to sounds and vocal gestures and which are communicated from mother to child? Or could it be that something of this “sorrowing” was immanent in the sounds themselves and their relations?

The central problem addressed by J. Shephard and Peter Wicke in *Music and Cultural Theory* concerns what they see as a pervasive tendency within cultural studies to elide the role of music as sound in the construction of meaningful cultural practice and to treat musical forms as mere ciphers for culture to inscribe its projects. On the other hand, the authors reject the opposite tendency: that of treating music as a self-contained and self-referring world of sound that is somehow immanently meaningful in itself. Shephard and Wicke argue that attempts by cultural theorists and musicologists to theorize music as a cultural practice are all somehow flawed; in their different ways, they fail to provide an adequate conceptualization of music as a constitutive social practice, one which does equal justice to its social constitution as well as to the irreducible contribution made by the sound material of music.

The greater part of the book is devoted to pointing out this inadequacy in a survey of theories about music and signification. The authors dip into the ideas of theorists ranging from Leonard Meyer, who writes on musical meaning and affect, to the work of French linguistic thinkers. All the usual suspects are here from Saussure, Barthes, Kristeva, Derrida through to Lacanian psychoanalytical contributions and to the contributions of Tagg, Middleton, and a host of others. The method is invariably the same in each case. Each theorist can apparently take us some of the way to the philosopher’s stone but ultimately fails at a key point and another theorist’s work is immediately engaged. There are repeated intimations of a great theoretical leap to be taken in the next paragraph, page, chapter, and so on of a conceptual landmark that is just about to be reached. When we get there, we find only the authors armed with spades and up to their armpits in a conceptual hole that continues to grow larger the more they dig.

There is certainly something for the reader to gain from being introduced—albeit idiosyncratically—to the ideas of a great many cultural theorists. There are also moments when the authors’ interrogation of the ideas of others appears to be getting somewhere—particularly with respect to the discussions of the theoretical importance of the concept of homology—and the reader’s interest is then awakened once more. However, the ponderous theorizing of the authors continually intrudes and prevents these excursions into the literature from being a useful or adequate survey of the field.

The book is dense, jargonistic, and difficult to read. And yet, for all the material used, there is one glaring absence—so loud that it screams. Theodor Adorno has arguably had more to say about music, language, culture, and society than almost any of the theorists discussed by Shephard and Wicke and, more than them, has written extensively about music in a way that respects the contribution of sound as material and music

as social praxis. His work is discussed nowhere in this book. His name does not even appear in the index and is mentioned only once in a reference in the text.

All the difficulties of this book as a reading experience as well as a theoretical document are best conveyed through a passage on page 162, which is one of those trumpeted theoretical climaxes through which—we are led to believe—the scales will fall from our eyes:

It is at this point that the real difference between language and music as structures of articulation becomes clear. The difference lies not in a difference between "difference" and "equivalence," or in a difference between "opposition" and "similarity" of elements of signification. It lies within a different order of differences. Difference in language is based on a principle of repulsion which results in the order of relations obtaining between elements of signification deploying components of meaning in an exclusive, consistent and, as a consequence, "immanent" association with sounds recognised as discrete. Difference in music is based on a principle of attraction which results in the order of relations obtaining between similar and different elements of signification deploying components of meaning throughout concatenations of sound recognized as discrete in a manner which has the potential to become infinitely complex. It is this basic difference in the character of how music and language function as structures which ultimately clarifies the somewhat confusing use of terminology in the discussion of these issues.

I am reminded of Adorno's observation that Wagner's music is full of claims of brave deeds to be done, of battles to be won, and enemies to be slain. These are all mendacious, he suggests, because there are no real enemies to be found within his music to subdue. Perhaps the authors of this book might have succeeded in convincing this reader of the reality of their theoretical battles, of the enemies to be slain in their text, had they managed to deliver what is promised in their critical approach, namely an idea of their own that is sufficiently clear and distinctive, as truly to cast light upon the errors of others. In the introduction, they promise us that we will see that in chapter 9. Alas, I searched for it there and in all the other chapters but was unable to find it.

The Rules of Art: Genesis and Structure of the Literary Field. By Pierre Bourdieu. Translated by Susan Emanuel. Stanford, Calif.: Stanford University Press, 1996. Pp. xx+408. \$49.50 (cloth); \$18.95 (paper).

Wendy Griswold
Northwestern University

So smart. So irritating. The smart part is the way in which Pierre Bourdieu has combined his theories—symbolic versus economic capital, fields, the habitus—with his detailed knowledge of 19th-century French literature to illuminate both the literary system and the literary texts.

Focusing on Flaubert, Bourdieu sees him engaged in a double refusal, both of the commercial success encouraged by the fiction market and of the social engagement urged by bohemian aesthetes. While Flaubert's position is well known, Bourdieu argues that it can best be understood through his location within a particular literary field, that of mid-19th-century Paris, where young writers attempted to make their careers among the contradictory demands of art and power (economic and political). From this follows a detailed analysis of the possibilities open to a writer in this location, especially whether to embrace a bohemian lifestyle and radical politics or to write for the market.

Bourdieu insists that a writer makes choices and is an active agent, not a mere conduit for social forces (here he takes exception to Lucien Goldmann's genetic structuralism and to Foucault's discourse theory), but that the alternatives open—the "space of the possibles"—are limited by the writer's field position. Some of the most important choices a writer makes shape the literary texts themselves. Bourdieu offers a brilliant reading of *A Sentimental Education* to show how Frédéric's indeterminacy, his inability to decide between the worlds of art and politics, is homologous to Flaubert's own understanding of the author's dilemma, and that this dilemma itself is the product of literature's successful assertion of autonomy in a world organized by commercial markets. Flaubert, along with Baudelaire, was a major voice in asserting this autonomy—art for art's sake—but at the same time in *A Sentimental Education* he coolly diagnosed the paralysis it produced in weaker souls like Frédéric.

The literary field organized in the 19th century is essentially intact today. Bourdieu sees writers, and other artists and intellectuals, still torn between the opposing economic logics of symbolic capital and economic capital. The two are inversely related; for example, a writer who attains commercial success thereby forfeits his status in the zone of pure art. The hard won autonomy of literature and the arts is useful politically, however; Zola's intervention in the Dreyfus affair, for example, carried moral authority precisely because of the presupposition of the disinterested man of letters. (In a postscript Bourdieu urges contemporary intellectuals to defend their autonomy.)

Bourdieu has some wise things to say about method. He denies that an internal, formalist analysis (a close reading of the text and only the text) and an external, sociological one (the text as dependent variable) are irreconcilable. Instead, there is "a homology between the space of works defined by their essentially symbolic content . . . and the space of positions in the field of production" (p. 205). This is the case because that set of generative dispositions that Bourdieu calls the *habitus* is produced and reproduced by the writer's position in the literary field (and that field's position in the field of power), and this *habitus* will itself structure the decisions a writer makes in forming the text. Better writers make better choices and execute them more skillfully, but the *habitus* that shapes their artistic creations is the product of particular historical conflicts. It is the *habitus* of the discerning reader, as well.

Now for the irritations. To describe the organization of the book as haphazard is to be charitable. It is as if Bourdieu cleaned out his desk and stuck a staple through everything that involved literature. I suggest that the reader should start with the beginning of part 2, "Questions of Method," then read part 1 on the field of 19th-century French literature, and then read the Flaubert prologue to see what difference Bourdieu's theory means for the interpretation of the text. Part 3 on readers is not as innovative, although Bourdieu shows how art historian Michael Baxandall's period eye is essentially the same as the habitus, and the last chapter of part 2, "The Author's Point of View," is something of a hodgepodge.

Even more irritating is Bourdieu's reluctance to relate his work to that of others. A great deal of his theory sounds like Lucien Goldmann's genetic structuralism, and it hardly seems sufficient to say that Goldmann did not see authors as conscious agents and Bourdieu does. The theory of fields would benefit by drawing on network analyses of Karl Rosengren and Helmut Anheier, the relations among participants in the literary system could stand some grounding in the institutionalist analysis of Paul DiMaggio or the Dutch research group including Cees van Rees and the Poetics circle, the discussion of autonomy would benefit by comparing it to Peter Bürger's account of the avant-garde, and the consideration of the power/culture relationship would surely be stronger if insights of Raymond Williams and Terry Eagleton on the one hand, and post-structuralists on the other (Foucault is barely mentioned, and Derrida not at all) were taken into account, even in refutation. The point here is not that the book contains insufficient citations—"Some people read the literature; I write the literature" was what we used to say in graduate school, and no one has made a more important contribution to cultural sociology than Pierre Bourdieu—but that the sociology of literature is cumulative. A lot of theoretical and empirical work has been going on in recent years, and either Bourdieu is wasting some of his energies by reinventing the wheel or he is not giving sufficient credit to the other wheelwrights.

A third irritation is the book's narrowness—it is just too French! Decades ago Cesar Graña (another scholar Bourdieu does not cite) described the tensions of 19th-century French writers, drawn to Paris, fashioning a bohemian lifestyle in which the necessities prompted by financial hardship became virtues, and increasingly dependent on a bourgeois market that they increasingly despised. This created an unusual breach between cultural and economic capital. Bourdieu argues that there is an inverse relationship between prestige within the subculture of writers and commercial success, but is this more true of France, the result of this 19th-century legacy, then elsewhere? Since Bourdieu makes no effort to look at literature or literary systems that are not French, he cannot answer this question. It may well be that fields are messier elsewhere, or are constructed (as Michèle Lamont has suggested) around different poles. Eschewing any comparative analysis has the effect of making Bourdieu's

theory seem more limited than it may actually be. So irritating. But so smart.

The Business of Children's Entertainment. By Norma Odom Pecora. New York: Guilford Press, 1998. Pp. 190. \$30.00.

Heather Hendershot
Queens College, CUNY

Although Disney does not tolerate copyright infringement, most creative children's media researchers find a way to represent Mickey Mouse. The cover of Eric Smoodin's *Disney Discourse* (Routledge, 1994) features ghostly hand-drawn Mickey mug shots; *From Mouse to Mermaid* (Elizabeth Bell et al., eds. [Indiana University Press, 1995]) shows only the elusive rodent's ears. Norma Odom Pecora's *The Business of Children's Entertainment* pictures a gloved, four-fingered hand clutching a bag of money. Pointedly highlighting the profit motive behind children's media, Pecora's cover recalls the flowing gold coins on the front of Ariel Dorfman and Armand Mattelart's *How to Read Donald Duck* (I.G. Editions, 1995), which offered the first scholarly assessment of Disney's profitable imperialism. Pecora's resemblance to previous children's media researchers ends with her cover, however. While most research on children's media grapples with ideological issues, Pecora, with relentless objectivity, reveals the nitty-gritty economics of the children's market. She gives comprehensive data on films, videos, records, and books but mainly targets television.

The primary sources for this exhaustively researched book are trades such as *Advertising Age*, *Variety*, *Broadcasting and Cable*, and *Playthings*. Pecora also uses the only previous economic study of children's media, William Melody's *Children's Television: The Economics of Exploitation* (Yale University Press, 1973). Additionally, Pecora has generated original data: a quantitative survey of media research on children's advertising (compiled with Ellen Wartella) and, most impressively, a *TV Guide* database used throughout the book. While such data will prove useful to other researchers, Pecora's methodology will frustrate historians. She never questions her sources. A strictly economic study, her book cannot address the irrational or stomach contradictions.

Pecora is strongest when she narrows her data and performs discrete case studies. Chapter 4 dissects the Smurfs, He-Man, and ThunderCats, all characters linked to TV shows and licensed products. Pecora traces changes in contractual alliances in the toy and TV business over the course of the eighties, arguing that these three cases illustrate a progressive consolidation of control of all aspects of TV and licensing under fewer and fewer companies. In chapter 5, Pecora scrutinizes nonnetwork TV outlets. This chapter constitutes the only research to date on Nickel-

odeon, the cable venue with the highest rated children's programs. Nickelodeon started without commercials but quickly succumbed to the market. The Disney Channel, conversely, is ad-free, but—surprise—the whole channel is a plug for the Disney brand. PBS is no less commercial, as Pecora explains in her thorough account of the rapidly increasing commercialization of children's PBS programs.

Other chapters of Pecora's book amass information but are less compelling. Chapter 1 describes how the child's consumer status has shifted since the 1920s. Radio advertisers realized children could influence parental decisions; by the 1980s, corporations (IBM, Sony) viewed children as sophisticated consumers in their own right and realized they could target kids and inculcate lifelong brand loyalty. Chapter 2 summarizes historical changes in children's TV sponsorship patterns and regulatory policy. Chapter 3 explains TV barter and syndication and toy marketing and licensing. Here, Pecora's explanation of how the FCC's prime time access rules affected children's programming is unique. Indeed, perhaps the most valuable aspect of the book is Pecora's understanding of the place the children's market occupies within the broader context of regulation and the entire television industry. Chapter 6 briefly surveys the book, magazine, recording, video game, film, and home video markets, and chapter 7 explains the internationalization of children's media, focusing on Saban Entertainment (creators of *Power Rangers*) and Zodiak, a U.S.-based company (now defunct) financed by foreign investors and specializing in characters created specifically for the global market. Pecora agrees with Haim Saban: the world is "one big boundaryless marketplace" (p. 150). This chapter will frustrate readers hungry for a more critical appraisal of the global marketplace, yet information compiled here will also prove useful to such readers.

It is a relief to read a book that explains the economics of children's media without sentimentalizing children as "TV victims" or imploring readers to kill their televisions, but ultimately, this book is frustrating in its lack of passion. Pecora does not advocate TV celibacy, TV activism, or regulatory protectionism nor does she address the child's potentially empowering use of TV, as the sociologically oriented Patricia Palmer, Ellen Seiter, and David Buckingham have done. Instead, she is relentlessly descriptive. Her brief conclusion laments that U.S. children are "fair game" for marketers, but, she argues, "short of revolution, it would appear there is little we can do" (p. 158). Perhaps children's culture can never be decommercialized, but it is simply untrue that there is nothing to be done. Media literacy instructors teach children to take cameras into their own hands, while activists attack the public school presence of the advertising driven Channel One. The Children's Television Act of 1990, the product of Action for Children's Television's relentless activism, does not challenge the entertainment industry's structure, but it does force broadcasters to change their programming in hopes of both placating the FCC and disarming activists. What the profit-driven industry that Pecora describes does not realize is that just as crafty publishers will always find

ways to represent Mickey on book covers, media activists will never be fully defused by the industry's bomb squads.

Community Organizing for Urban School Reform. By Dennis Shirley. Austin: University of Texas Press, 1997. Pp. x+338. \$35.00 (cloth); \$17.95 (paper).

Jean Anyon
Rutgers University

A national consensus has emerged that several decades of school reform activity have failed to render public schools in U.S. low-income urban neighborhoods successful centers of academic excellence. Rather, inner-city schools are typically underfinanced, dreary institutions, too often staffed with political appointees having little preparation in teaching or administration. Students are too often taught little that is edifying or useful. Despite almost 30 years of effort by school reform groups, the situation remains very much the same as described in the 1960s and 1970s by Kenneth Clark, Herbert Kohl, and Jonathan Kozol.

I argue in my new book, *Ghetto Schooling: A Political Economy of Urban Educational Reform* (Teachers College Press, 1997), that school reform has failed in our cities in large part because reformers and politicians have ignored the powerful effects on schooling and school reform of the poverty and political and economic isolation of inner-city residents. Isolation of low-income residents from jobs, decent housing, health care, and political representation has a devastating effect on attempts to improve inner-city schools.

This book by Dennis Shirley describes an important strategy for urban school improvement that has been ignored by school reformers. The strategy is community organizing—exemplified in the book by a group called the Industrial Areas Foundation (IAF). The goal of this group is helping low-income urban residents develop their abilities to change their own neighborhoods and local institutions. The IAF is descended from the group of the same name founded by the late Saul Alinsky in 1940. Ernesto Cortes, one of the current leaders, received a MacArthur “genius award” and was highlighted by Bill Moyers on the PBS television series, “A World of Ideas.”

Current IAF community organizing tactics are more nuanced in political rhetoric and strategy than those of the flamboyant and confrontational Alinsky. IAF currently provides leadership training for over 30 organizations representing nearly 1,000 institutions and over 1 million families. It works through neighborhood churches and families to build civic engagement; IAF organizers teach community residents the social capital to resolve their grievances against the city and the school systems. These skills include how to speak in public, lead actions, take risks, and guide others; how to develop social and political relationships within which to

challenge the indifference and apathy of corporate and government officials; how to negotiate with the holders of power, compromise, confront when necessary, and rebuild collaboration.

IAF organizers begin in a community by working to improve visible, concrete problems, like blocked sewers or inadequate neighborhood police protection, in order to have a success that will make the residents feel comfortable with them and will make the city government take notice. Then, IAF organizers work with residents to help them take action to solve their own problems, particularly in regard to failing local schools.

The educational goals of the IAF are not merely to add curriculum or other typical reform programs to schools, although they support those. Their main educational goal is to transform the culture of the school by transforming the way parents view themselves. The IAF wants the community to see itself in a new way, as actors in the urban drama, not just as passive victims or helpless bystanders. Simply, they want to engage the community in neighborhood and school improvement. They have been very successful in cities in which they have worked—primarily in Texas and California, but also in Maryland, New York, and Tennessee. They have been funded by the Catholic Church and the Ford, Rockefeller, and Annenberg foundations.

Dennis Shirley vividly portrays the activities of the IAF. The first part of the book describes the origins of the group and the current context of inner-city poverty in which they work. Part two describes in detail how the IAF operates and how the group has developed civic capacity in parents and community residents—and resulting improvements parents instigated in schools—in three cities in Texas. The book's third part offers an analysis and critique of the IAF approach. Included in this critique is the wise advice that not only do educators need to learn more about community groups, but such groups—and the IAF as well—need to learn more about the professional world of educators. All who care about our inner-city schools need to collaborate. Shirley is fair, thorough, and thoughtful in his analysis and provides important food for thought for all those who would work to make inner cities, and their often desperate schools, decent places to live and learn.

Against the Multicultural Agenda: A Critical Thinking Alternative. By Yehudi O. Webster. Westport, Conn.: Praeger, 1997. Pp. 228. \$59.95 (cloth); \$22.95 (paper).

James A. Banks
University of Washington

Webster provides a critique of multicultural education and multiculturalism and argues that schools and universities should teach critical thinking rather than multicultural education. He views multicultural education and critical thinking as inconsistent. Webster views multicultural

education as "advocacy" rather than as an interdisciplinary field of study. Webster argues that schools must, in effect, be apolitical places that only focus on the process of critical thinking, and should not have normative goals, such as equity and human rights. He writes, "Critical thinking reforms are aimed at improving intellectual competencies of all those involved in learning, irrespective of whether this improvement leads to more productive workers, greater social harmony, a strengthening of democracy, socioeconomic equality, or the liberation of the oppressed" (p. 188).

Webster's book has several major flaws that seriously limit its usefulness. He criticizes multicultural educators for being advocates rather than dispassionate scholars, yet his book epitomizes advocacy, as its title indicates. Rather than provide an objective, scholarly discussion of the field, his book is a brief against multicultural education. Webster has a very confused conception of who the multiculturalists are. He describes scholars with widely divergent ideological views as multiculturalists, including scholars who oppose multicultural education such as Arthur Schlesinger and Lynne Cheney. He also describes Molefi Asante, the noted Afrocentrist, as a multiculturalist. Webster also describes Afrocentrism as a part of multicultural education, which is a serious mistake. The confusion he perceives among the multiculturalists results in large part from the inaccurate way in which he conceptualizes the field.

Webster's inability to understand that education inevitably has normative goals and cannot take place within an ideological vacuum is one of the most serious shortcomings of his book. Students cannot be taught to think without the use of teaching materials that have cultural assumptions and values. A key aim of multicultural education is to promote critical thought and action. Multicultural theorists assume that one of the most effective ways to teach students to critique their own assumptions and beliefs is to juxtapose them with those of people from other cultural groups. Consequently, multicultural education and critical thinking are highly compatible. Multicultural content helps students to uncover the cultural assumptions of content as well as their own cultural assumptions and biases, which greatly enhances their ability to think.

Children, Schools, and Inequality. Edited by Doris R. Entwisle, Karl L. Alexander, and Linda Steffel Olson. Boulder, Colo.: Westview Press, 1997. Pp. xiii+238. \$59.00

Barbara Heyns
New York University

This study integrates sociological research on status attainment and the life course with ideas drawn from ecological studies of child development. It is the first book to focus in-depth on early schooling and

preadolescence, based on the Baltimore Beginning School Study (BBS) at Johns Hopkins University. The authors have undoubtedly collected the most detailed and comprehensive longitudinal data on children's home environment and school outcomes, from school entry in 1982 through high school graduation. For sociologists of education, the book has few surprises. For researchers in general, however, this book represents an opportunity to catch up with a wealth of research on schools. This volume is the culmination of a decade of work and numerous articles, but the wealth of BBS data suggests more volumes to come in later years.

The authors make a very convincing case for the importance of educational transitions, especially into first grade. At every grade level, however, a shocking number of children fail to progress. Baltimore has apparently abandoned "social promotion" as school policy. By sixth grade, 43% of the children had been retained at least one grade. First graders are almost twice as likely to be retained as older children, but their problems do not end there; almost half of the first grade failures enter sixth grade two or more years behind their age peers. As one might guess, the unsuccessful first graders tend to be from single-parent families, with little education and few resources, who experience numerous economic difficulties and considerable stress during the study. These children are also likely to miss a lot of school, to be in the lowest reading group, to score substantially below grade level in achievement, to attend school in relatively poor neighborhoods, to have changed schools two or more times before fifth grade, and to receive the poorest marks at every grade, even those they repeat. Their parents and their teachers both expect them to achieve well below average.

The authors unravel competing explanations for early gender differences in achievement and retention. Teachers view boys as less cooperative, attentive, or interested than the girls throughout the early grades; consequently, boys get lower citizenship and reading scores than girls do. Part of this gender difference may in fact be the result of teacher favoritism. Both boys and girls do better with teachers who prefer their gender, although boys get more attention than girls, this is due to their deportment not their achievement. Interestingly, by the end of first grade, girls like school, their teachers, and reading better than the boys do, although they enter school with similar attitudes. These gender differences persist into adolescence.

Throughout the book, the authors detail the implications of their study for school effects and for educational policy, showing quite clearly that achievement is not necessarily linear with respect to time. It is gratifying to have one's own work replicated in such a thoroughly thoughtful and insightful manner—for younger students in a different urban area and over a longer period of time. During the school year, children from diverse backgrounds and neighborhoods—irrespective of age, race, or gender—gain achievement points at close to the same rate; hence, virtually all the observed inequality in outcomes cannot be attributed to schooling.

Socioeconomic gaps in achievement accrue during the summer, when children are not in school.

The authors propose a "faucet theory" to account for school effects in the patterns of cognitive development. During the school year, learning rates of poor children are almost as high as those of their more advantaged peers; thus, summer learning—or the unique effects of home background—accounts for 80–100% of aggregate annual inequality, despite the fact that most achievement gains are attributed to the school year. The "faucet theory" assumes that schooling is turned on and off according to the season. School-year gains result from the flow of intellectual stimulation found in school; during the summer, learning rates wither and fall because poor families and poor neighborhoods can only provide a trickle of the mental dousing children get in school.

While this book replicates previous findings on summertime learning, it avoids the most intractable questions about children's progress in school. Demonstrating that test scores are not linear with respect to time begs the question of whether gains are equivalent at different points on the scale. The truth is we have no idea how much children have learned when their test scores increase—or decrease—by a certain number of points. Achievement tests are ordinal measures, while an analysis of gains requires imputing or assuming equal intervals. For example, children scoring below the mean tend to lose points during the summer. We cannot prove, however, that gains and losses are equivalent. If upper-status children improve at a faster rate during the summer months than poor children, this could mean that their parents are more effective at promoting either learning or retention. It could also mean that standardized tests constructed for successive years are normed to give meaningful intervals only during the school year. Hence, one finds equal school year gains because tests were constructed and normed to yield linear increases with respect to age or months in school. Uneven summer growth—or differential "forgetting"—cannot be proven without interval measurement. Until this problem is solved, sociological theories of learning cannot refute the conventional—and unpalatable—wisdom supported by psychologists: test scores are essentially stable after a child enters grade school. Until we have measures of learning and not just an ordinal ranking, sociologists of education are rolling a very large boulder uphill.

Motivation and Delinquency. Vol. 44 of the Nebraska Symposium on Motivation. Edited by D. Wayne Osgood. Lincoln: University of Nebraska Press, 1997. Pp. xiii+295. \$40.00.

Simon I. Singer
University at Buffalo—SUNY

Readers looking for a singular, coherent explanation of motivation and delinquency will not find it in this edited book of five articles. Instead,

they will discover a diverse set of research perspectives drawn together by Wayne Osgood as part of the distinguished Nebraska series on motivation. In the introduction, Osgood makes note of the fact that the public's perception of juvenile crime is "often out of touch with the evidence" (p. ix). Despite this introductory concern with juvenile crime, a considerable amount of the material is concerned with behavior that is defined in various ways as antisocial or deviant.

The first article, by Joan McCord, draws on the ancient Greek philosophers to address some major criminological assumptions. She critiques rational choice and social control theories for their "egoistic" assumptions and for not taking human altruism as a serious possibility. For that critique alone, I consider the article worthy of study. Yet McCord manages in just 10 concluding pages to go a step further to present a "construct theory of motivation" and to relate it to delinquency via "potentiating reasons." However, in her too brief a "sketch of a theory," it is difficult to see how potentiating reasons differs from techniques of neutralization or concepts that relate a juveniles justification of past and future acts of delinquency.

The next article, by Michael Rutter and colleagues, presents several data sets from their long-term research program to make the case for the "heterogeneity of antisocial behavior." That there is heterogeneity should not come to any surprise to anyone who has studied delinquency. But the diverse ways in which problem behaviors are correlated in the numerous data sets that this article draws on is quite bewildering, particularly for those of us who are unfamiliar with the technical definitions of some forms of antisocial behavior. There are some unexpected findings, such as hyperactivity being related to adult crime but only for those without reading difficulties (p. 66). Why that is the case is hardly discussed, although sociologists would appreciate part of the conclusion, which states that "motivation stems from the emotions and thought processes of an individual in a particular social context" (p. 107).

The article by Gerald Patterson and Karen Yoerger relates child-rearing practices to early onset delinquencies. As might be expected, the model that they generate is behavioral and stresses the functional aspects of antisocial behavior within families. The behavioristic approach places kids on automatic pilot in their learning of antisocial ways to resolve disputes. They also make the point that antisocial behavior is more a product of poor parenting than is the case for late onset delinquents. Late onset delinquents produce more covert acts of delinquency and are less serious delinquents than early onset delinquents. As in the previous article, delinquency is too readily associated with conduct that is subjectively evaluated by nonlegal officials, such as parents and teachers.

The article by Tedeschi is of a different stroke. The emphasis is on social interaction and how individuals respond to perceived threats and injustices. Something more cognitive is going on than the behavioristic approach that Patterson and Yoerger present. The cognitive interacts with the behavioral in that certain individuals are less capable of re-

sponding to perceived threats in a noncoercive manner. What needs to be explained is not the motivation for delinquency but the motivation for control. Coercive forms of behavior work for some individuals and not for others. It may be difficult to see the social in social interactionist theory if not for the sense of justice that Tedeschi brings into his discussion.

The final article, by Heimer and Matsueda, in my mind brings it all together. They make the case for symbolic interactionism with its emphasis on the "primacy of the group or society." Of particular relevance to the developmental aspects of life course delinquency research is their point that cognitive aspects are more likely to work in problematic situations than in nonproblematic situations. This takes into account variation in life course patterns that at different stages would bring into play the very diverse ways in which meaning and behavior account for deviance. The authors review a considerable amount of delinquency theory and relate how symbolic interactionism fits in. Moreover, as other authors do, they are able to draw on a variety of research findings that tend to support the symbolic interactionist perspective.

Although each of the articles can be commended for their sensitivity to life course issues and for discussing fundamental issues in psychologically driven delinquency research, they lose sight of an important distinction in the term "delinquency." Delinquency is actually more serious than antisocial behavior, hyperactivity, conduct disorder, or deviance. It has unfortunately been relegated to something less serious than what the public generally defines as juvenile crime. Psychological and sociological self-report ratings have trivialized what amounts to an official response and a particular legal definition of behavior. An understanding of the motivation for delinquency would be more complete if we consider how officials in a variety of official capacities come to define kids as delinquent. Despite the neglect of this crucial aspect of delinquency research, the articles in this book do go a long way in helping us to understand the individual psychologies motivating a considerable amount of delinquent behavior.